UNIVERSIDAD AUTÓNOMA DE MADRID
ESCUELA POLITÉCNICA SUPERIOR

TRABAJO FIN DE MÁSTER

Reconocimiento facial no invasivo en dispositivos móviles

Máster Universitario en Ingeniería de Telecomunicación

Autor: David Costa Da Silva
Tutor: Belén Ríos Sánchez
Ponente: Julián Fiérrez Aguilar

FECHA: Febrero 2018

Resumen

Con la expansión de la tecnología, y particularmente el uso de los teléfonos móviles, son cada vez más las aplicaciones y servicios que necesitan autenticar al usuario. Para ello, el mecanismo de seguridad más utilizado hoy en día son las contraseñas, cuya vulnerabilidad hace presente la necesidad de buscar nuevos métodos más seguros y eficaces. Así, la autenticación biométrica de los usuarios mediante lo que ellos son y no solamente lo que saben, se presenta como una alternativa eficaz apoyada en los resultados de varias décadas de investigación en el reconocimiento de rasgos como la huella dactilar, el iris o la cara.

De este modo, en los últimos años la biometría ha pasado a ser una parte fundamental en nuestras vidas. Pese a que la biometría era una ciencia relativamente desconocida para la mayoría de la población hace unos años, las múltiples investigaciones en este campo durante las últimas décadas, junto con su integración en la tecnología de última generación, han hecho que la gente comience a comprender su gran potencial en la actualidad y en el futuro. De este modo, lo que hace unos años se consideraba ciencia ficción se ha vuelto una realidad, y hoy en día el desbloqueo de ciertos dispositivos tecnológicos utilizando alguna parte de nuestro cuerpo es un gesto habitual.

Por otro lado, el elevado número de autenticaciones requeridas a lo largo del día por dichos servicios y aplicaciones ha puesto de manifiesto la necesidad de realizarlas de la forma más natural posible. Así, en los últimos años ha cobrado fuerza la identificación no invasiva, en la que se potencia la comodidad del usuario manteniendo los niveles de seguridad.

Además, el uso de dispositivos móviles presenta nuevos retos en la identificación biométrica. Estos nuevos desafíos son originados principalmente por la variabilidad de las condiciones de captura, consecuencia directa de la variedad de ambientes en los que estos dispositivos son utilizados.

Este Trabajo de Fin de Máster se centra en el estudio y análisis de técnicas tradicionales para el reconocimiento facial no intrusivo orientado a dispositivos móviles. En concreto en este trabajo se estudian 3 técnicas diferentes, una aproximación geométrica que extrae la información geométrica derivada de una serie de puntos característicos de la cara, y dos aproximaciones holísticas que utilizan información de la textura de la imagen para el reconocimiento de los usuarios. Tras el estudio individual de estas técnicas, se realizará una fusión de los datos proporcionados por las técnicas geométrica y holśitica con el fin de evaluar si la combinación de información repercute en una mejora del sistema.

Palabras Clave

Biometría, biometría no intrusiva, puntos característicos faciales, patrones binarios locales (LBP), patrones derivativos locales (LDP), dispositivos móviles, landmarks, verificación.

Abstract

In the new era of the technology, there are many applications and services that need to authenticate the users. Nowadays, passwords are the most common and security mechanism used to this end, but their vulnerability reveals the need to search for new, safer and more efficient methods. Biometric authentication of the users through what they are and not only what they know is an effective alternative supported by the results of several decades of research in the field using traits such as fingerprint, the iris or the face.

Accordingly, in recent years, biometrics has become a fundamental part of our lives. The large researches in this field during the last decade and people interaction with this technology, made that people begin to understand the great potential of biometrics. In this way, what was considered science fiction a few years ago, has now become a reality, and today the unlocking of certain technological devices using a part of our body is a usual gesture.

On the other hand, the increasing number of authentications required along the day by this applications and services has revealed the need to perform them in the most possible natural way. Thus, in recent years, non-invasive identifications, in which user comfort is enhanced with safety has gained popularity.

In addition, the use of mobile devices presents new challenges in biometric identification. These new challenges are mainly caused by the variability of capture conditions.

This Master's Thesis is focussed on the study and analysis of traditional techniques for non-intrusive facial recognition in mobile devices. The biometric recognition techniques from facial images can be divided into two large groups: geometric techniques and texture-based techniques. The first ones are based on the genetic information derived from a series of characteristic points of the face, while the second ones use the texture of the image for the recognition of the users. Talking about this Project, three different techniques are studied, one based on the facial geometry and the other two on the texture of the image. After studying these techniques, a fusion of both methods will be done to evaluate if the combination of information can report better results in the biometric system.

Key words

Biometrics, non-intrusive biometrics, characteristic landmarks, local binary patterns (LBP), local derivative patterns (LDP), mobile devices, verification.

Agradecimientos

En primer lugar, quiero agradecerle a mi tutora, Belén Ríos Sánchez, su gran esfuerzo y dedicación en mi proyecto y todo lo que me ha ayudado para poder terminar con éxito este TFM.

También me gustaría agradecer a Carmen Sánchez Ávila y a mis compañeros del Grupo de Biometría, Bioseñales, Seguridad y Smart Mobility (GB2S) del Centro de Domótica Integral (CeDInt) la oportunidad de poder trabajar en el campo de la investigación biométrica y aprender con ellos estos meses.

Agradecer además a los profesores que me han enseñado y de los que he aprendido tantas cosas para pode realizar este proyecto.

También quiero agradecer principalmente el apoyo que me han mostrado mis padres, no solo en el ámbito académico, sino también a lo largo de toda mi vida. Ellos han hecho posible que haya podido continuar y finalizar mis estudios.

A mis compañeros de clase que han hecho que este año y medio de Máster haya sido más divertido de lo que habría sido sin ellos y que algunos de ellos sean ya mis amigos, además de ayudarme en muchos momentos del Máster.

Sin olvidarme de mis amigos, agradecerles su apoyo en mi vida en general y en los buenos momentos que hemos pasado todos estos años.

Finalmente, agradecerle a Emma estar ahí cuando más la necesitaba, intentar ayudarme en todo lo que ha podido y haber confiado en mí en todo momento.

Índice general

Índice de figuras

Índice de tablas

1

Introducción

La expansión de la tecnología, y en particular de los teléfonos móviles, ha derivado en una gran cantidad de aplicaciones y servicios que necesitan autenticar al usuario. Para ello, hoy en día el mecanismo más utilizado son las contraseñas, cuya vulnerabilidad hace presente la necesidad de buscar nuevos métodos más seguros y eficaces [1, 2]. Así, la autenticación biométrica de los usuarios mediante lo que ellos son y no solamente lo que saben, se presenta como una alternativa eficaz apoyada en los resultados de varias décadas de investigación en el reconocimiento de rasgos como la huella dactilar, el iris o la cara [3].

De este modo, la biometría ha pasado de ser una ciencia relativamente desconocida para la mayoría de la población a convertirse en una parte fundamental en nuestras vidas en los últimos años. Este cambio hsa sido principalmente motivado por las múltiples investigaciones en este campo durante las últimas décadas y su integración en la tecnología de última generación. De este modo, lo que hace unos años se consideraba ciencia ficción se ha vuelto una realidad, y hoy en día el desbloqueo de ciertos dispositivos tecnológicos utilizando alguna parte de nuestro cuerpo es un gesto habitual.

Por otro lado, el elevado número de autenticaciones requeridas a lo largo del día por dichos servicios y aplicaciones ha puesto de manifiesto la necesidad de autenticar a los usuarios de la forma más natural posible. Así, en los últimos años ha cobrado fuerza el reconocimiento biométrico no invasivo, que potencia la comodidad del usuario sin comprometer la seguridad [4]. Además, el uso de dispositivos móviles presenta nuevos retos en la identificación biométrica, originados principalmente por la variabilidad de las condiciones de captura, que es a su vez consecuencia directa de la variedad de ambientes en los que estos dispositivos son utilizados [5].

La biometría se basa en el hecho de que cada persona es única y diferente del resto de seres humanos. De este modo, la biometría estudia los rasgos físicos y conductuales que nos hacen únicos para crear sistemas de reconocimiento automático cada vez más seguros y robustos. Así, la biometría puede ser entendida como la aplicación de un análisis estadístico sobre datos biológicos.

Para poder ser considerada como un identificador biométrico una característica debe cumplir una serie de propiedades:

- Permanencia: la característica no debe cambiar a lo largo de la vida del individuo.
- Singularidad: la característica debe identificar al individuo de forma única.

- Facilidad de captura: la característica debe ser fácilmente adquirida, preferiblemente con equipo de bajo coste y de forma inmediata.
- Universalidad: la característica debe estar presente en todos los individuos. Colectividad: la característica debe poder ser medida de forma cuantitativa.
- Calidad: la característica debe aportar precisión, velocidad y robustez al sistema.
- Aceptabilidad: la característica debe contar con la aprobación por parte de los usuarios.

En la Tabla 1.1 se muestra una comparativa de las tecnologías biométricas más usadas en la actualidad. En ella se comparan características como la precisión, aceptabilidad, coste, facilidad de uso y facilidad de suplantación.

Tabla 1.1: Tecnologías biométricas más usadas en la actualidad

Tipo de biometría	**Huella digital**	**Geometría de mano**	**Iris**	**Voz**	**Cara**
Precisión	Alta	Media	Alta	Baja	Media
Aceptabilidad	Media	Media	Media	Alta	Media
Coste	Bajo	Alto	Alto	Bajo	Bajo
Facilidad de uso	Alta	Alta	Media	Alto	Media
Intrusismo	Bajo	No	No	No	No

Por regla general, las características físicas además de ser únicas en cada individuo, permanecen inalterables durante su vida, mientras que las conductuales pueden variar dependiendo de factores externos. Es por ello que las técnicas más usadas hoy en día en el reconocimiento biométrico en dispositivos móviles y plataformas que requieren una seguridad elevada están basadas en las características físicas del usuario.

Durante el funcionamiento típico de un sistema de estas características, el primer paso es dar de alta a los usuarios en el sistema. Para ello se capturan sus rasgos biométricos, se extraen las características que contienen y se almacenan en la base de datos para futuras comparaciones. Una vez introducida en el sistema información sobre los usuarios, se puede proceder al reconocimiento de los mismos. En la etapa de reconocimiento, los sistemas biométricos pueden funcionar en dos modos diferentes: verificación e identificación. En el caso de verificación, el usuario proporciona una identidad y el sistema comprueba su veracidad. Para ello, se captura y procesa una nueva muestra biométrica y se extrae el vector de características, que se compara con el patrón almacenado en la base de datos. De este modo se realiza una comparación 1:1 entre la nueva muestra y el patrón almacenado. La comparación puede variar dependiendo de la política adoptada por el sistema, pero generalmente devuelve un valor numérico que muestra el grado de similitud entre la muestra capturada y el patrón de ese usuario. Este valor es recogido por el módulo de decisión, que lo compara con un umbral previamente establecido para confirmar o negar la identidad del usuario. En el caso de identificación, el usuario simplemente proporciona su muestra biométrica y el sistema busca entre los usuarios que tiene almacenados aquel (o aquellos) cuyo patrón es más similar a dicha muestra. En este caso la nueva muestra capturada se compara con todos los patrones almacenados en el sistema. Es decir, se realiza una comparación 1:N, siendo N el número total de individuos enrolados en el sistema. El resultado de esta comparación será una lista ordenada de candidatos según el grado de similitud entre el

nuevo vector de características y los patrones almacenados. La salida del módulo de decisión será el usuario o usuarios con mayor similitud.

La biometría facial permite reconocer a los usuarios analizando imágenes de su cara. A diferencia de otras técnicas biométricas como la huella dactilar, la voz o el iris. Esta tecnología no es intrusiva, y no requiere necesariamente la colaboración del usuario. De hecho, las características podrían ser adquiridas sin que el usuario advierta que está siendo grabado, al igual que ocurre en ciertos sistemas de videovigilancia y CCTV. Además, es una técnica que goza de buena aceptación por parte de los usuarios, ya que el rostro es el principal rasgo que los seres humanos utilizamos para reconocernos.

Las técnicas de reconocimiento biométrico a partir de imágenes faciales se pueden dividir en dos grandes grupos: técnicas geométricas y técnicas holísticas. Las primeras están basadas en la información geométrica derivada de una serie de puntos característicos de la cara, mientras que las segundas utilizan información de la textura de la imagen para el reconocimiento de los usuarios. En concreto, en este trabajo se estudian tres técnicas diferentes, una basada en la geometría facial y las dos restantes basadas en la textura de la imagen. En el primer caso, a partir de una serie de puntos característicos de la cara, también conocidos como *Landmarks*, se extraen una serie de características geométricas tales como distancias y/o ángulos. En esta aproximación, la detección precisa de dichos puntos es extremadamente relevante, por lo que se han estudiado varias librerías de detección de puntos faciales con el fin de analizar y comparar los resultados finales obtenidos con cada una de ellas. En los otros dos casos se utilizan descriptores de textura para extraer la información biométrica contenida en las imágenes: patrones binarios locales y patrones derivativos locales. Ambos son descriptores bien conocidos que han sido aplicados con anterioridad a la biometría facial. El fin de su estudio en este trabajo es comprobar su aplicabilidad a imágenes no intrusivas captadas desde teléfonos móviles.

Dada su orientación a dispositivos móviles, el sistema propuesto en este TFM es un sistema biométrico de verificación. Su estructura sigue el esquema típico de estas características (Figura 1.1). Así, el sistema biométrico comprueba si el usuario que está intentando acceder al sistema es el propietario del dispositivo o no, con el objetivo de garantizar la protección de los datos y servicios ubicados en el mismo.

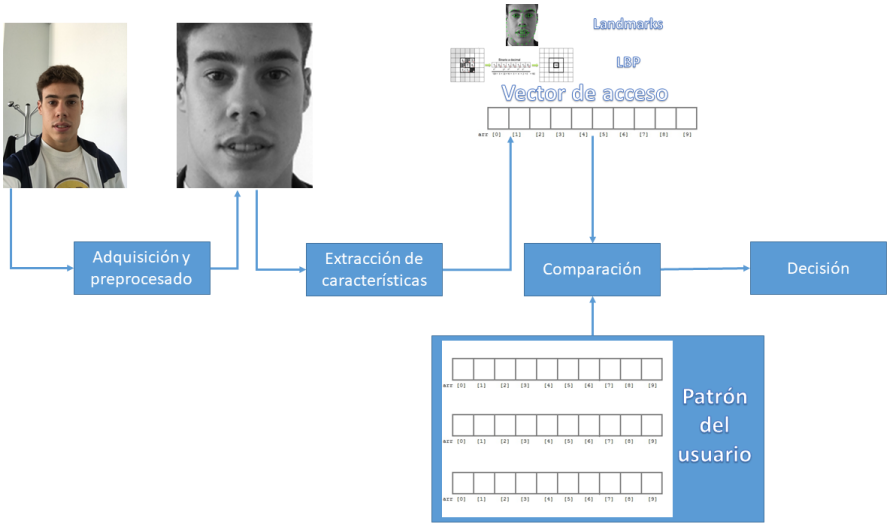

Figura 1.1: Sistema biométrico de verificación propuesto

El resto de esta memoria está organizada de la siguiente manera:

- En el capítulo 2 se expondrá el estado del arte, que se centrará en el análisis de trabajos previos de la biometría facial.
- En el siguiente capítulo se explicará la descripción del sistema. En él se detallarán los módulos de captura de datos y preprocesado, extracción de características, comparación, fusión y decisión.
- A continuación, se detallarán los experimentos y resultados obtenidos. También se describirán las bases de datos y los protocolos empleados para la evaluación del sistema.
- Finalmente, se expondrán las conclusiones obtenidas y se plantearán posibles trabajos futuros con los que continuar esta investigación.

2

Estado del arte

2.1. Introducción

En las últimas décadas la biometría facial se ha convertido en una técnica de reconocimiento muy popular debido a su aplicabilidad y facilidad de uso, ya que es posible integrarla en cualquier dispositivo que tenga una cámara. Las áreas de aplicación del reconocimiento facial son muchas y variadas, desde la seguridad en sistemas de video vigilancia por circuito cerrado de televisión, el control de fronteras o la investigación en bases de datos policiales, hasta la clasificación de género o el reconocimiento de expresiones y seguimiento de características faciales con el fin de monitorizar la fatiga o detectar stress [6]. Además, se trata de una técnica de gran aceptación por parte de los usuarios debido a su bajo coste, su carácter no intrusivo y el hecho de que el reconocimiento facial es una de las habilidades más significativas de la visión de humanos y primates. Es por ello que durante los últimos 20 años se han propuesto numerosas técnicas para el reconocimiento facial automático y se han desarrollado diversos estudios en relación a su rendimiento. Así, es posible encontrar en la literatura numerosos trabajos relacionados con técnicas generales empleadas para el reconocimiento facial [7, 8, 9].

Igualmente, con el objetivo de testear y evaluar los sistemas desarrollados, se han creado un gran número de bases de datos de imágenes faciales, muchas de ellas públicas, y se han diseñado diversos protocolos de evaluación que permiten comparar los resultados obtenidos entre distintas implementaciones. La Tabla 2.1 muestra una comparativa entre las bases de datos más utilizadas para la evaluación de sistemas de reconocimiento facial, mientras que en [10] se puede encontrar una recopilación exhaustiva y en continua actualización de las mismas. Por otro lado, en [11] se presenta una revisión de numerosas propuestas de evaluación sistemáticas de técnicas de reconocimiento facial, incluyendo los protocolos FERET [12, 13], FRVT 2000 [14], FRVT 2002 [15] y XM2VTS [16].

Tabla 2.1: Bases de datos para el reconocimiento facial

Nombre	**Año**	**Número de imágenes**	**Sujetos**	**Condiciones**	**Características del conjunto de datos**	**Referencia**
MOBIO	2010	193620	152	No controladas	Móvil	[17]
SecurePhone PDA	2006	12960	60	Controladas	Video, Móvil	[18]
FERET	1996	14051	1199	Semi-controladas	Luz, pose y tiempo	[12]
AR	1988	3288	116	Controladas	Luz y oclusión	[19]
CAS-PEAL	2003	30900	2747	Controladas	Luz, pose y accesorios	[20]
Face Recognition Data	1996	7900	395	No controladas	Etnia	[21]
SCFace	1996	41260	130	No controladas	Video	[22]
M2VTS	1998	N/A	295	Controladas	Tiempo, pose y multimodal	[16]
Yale B	2001	5850	10	Controladas	Pose y luz	[23]
CMU PIE	2000	41368	68	Controladas	Pose y luz	[24]
FIA	2004	12960	200	Controladas y no controladas	Pose	[25]
MIT-CBCL	1999	31022	100	Semi-controladas	Sin caras	[26]

No obstante, dadas las características de la biometría facial y el estado de madurez de la técnica, la actividad en este campo no se limita solamente a la investigación y son cada vez más los sistemas comerciales disponibles con este propósito: Viisage Technology [27], FaceKey Corp. [28], Cognitec Systems [29], ImageWare Sofware [30], BioID sensor fusion [31], Biometric Systems, Inc. [32], SpotIt for face composite [33]. Del mismo modo, también se ha producido un crecimiento de la aplicación de este tipo de técnicas en el ámbito digital, donde grandes empresas como Microsoft, Google, Facebook o Apple [34, 35], hacen uso de esta tecnología para analizar enormes cantidades de imágenes con el fin de identificar a las personas capturadas en ellas. Además, las aplicaciones móviles como AppLock desarrollada por Visidon para móviles Android [36, 37] están proliferando cada vez más.

2.2. Reconocimiento Facial

El reconocimiento facial se divide en dos tareas principales: la localización de la cara dentro de la imagen y el reconocimiento propiamente dicho, mediante el que la cara es asociada a un individuo previamente dado de alta en el sistema. Estas tareas se suelen llevar a cabo en tres pasos:

1. Detección de la cara dentro de la imagen. En este paso se evalúa si hay caras dentro de la imagen y en ese caso se devuelve su localización.

2. Extracción de las características faciales y composición del patrón biométrico. En este punto se busca una serie de características relevantes para la identificación del individuo que han sido previamente definidas. Dependiendo del método utilizado para la extracción de características, esta etapa puede requerir el entrenamiento previo del sistema.

3. Identificación del individuo mediante la comparación de patrones. Este paso está estrechamente ligado a la extracción de características y a menudo depende de un modelo generado a partir de dichas características.

La tarea de detección facial dentro de la imagen tiene un impacto directo sobre el resto del proceso, ya que debe afrontar varias dificultades como las variaciones en escala, localización, pose, expresión, condiciones de iluminación y/u oclusiones. En [38] se presenta un análisis de los avances recientes en detección facial, clasificando las distintas aproximaciones en cuatro categorías: métodos basados en el conocimiento, que utilizan técnicas predefinidas basadas en el conocimiento humano; métodos de características invariantes, que buscan las características de estructuras faciales robustas ante variaciones; métodos de comparación de patrones, que usan patrones de cara previamente calculados para decidir si hay una cara en una imagen, y métodos basados en apariencia, que aprenden modelos de un conjunto de imágenes de entrenamiento.

De acuerdo con [8], los métodos de extracción de características faciales pueden ser clasificados en dos categorías principales: patrones holísticos y patrones geométricos. Los primeros consideran la región facial completa como datos de entrada y obtienen una representación de la textura de la misma. Los segundos se basan en las características geométricas existentes entre ciertos puntos de interés relativos a los distintos elementos faciales (ojos, nariz, boca, etc.) también conocidos como *landmarks*. Estas características suelen ser principalmente distancias, posiciones relativas o ángulos. En [6] añade una tercera categoría a esta clasificación que engloba las soluciones híbridas formadas a partir de las dos anteriores. La gran mayoría de los sistemas actuales se engloban en la categoría de patrones holísticos, ya que son más fiables y su implementación es mucho más sencilla [39].

La mayoría de los trabajos realizados hasta la fecha se centran en los dos primeros pasos [7, 40, 41, 42], mientras que la comparación de patrones ha recibido menor atención. No obstante, en [43] se puede encontrar una evaluación de distintos clasificadores supervisados en términos de tasas de correcta identificación y tiempo de computación.

2.3. Reconocimiento Facial en Dispositivos Móviles

En la actualidad, con el uso cada vez más extendido de dispositivos móviles, la seguridad de los datos almacenados en ellos ha cobrado una gran relevancia. En este contexto, dada la mejora de las capacidades de este tipo de dispositivos, el reconocimiento facial permite al usuario evitar recordar códigos o contraseñas, proporcionando una mayor y más flexible seguridad. Así, es posible encontrar desde aplicaciones bancarias [44, 45] o plataformas de pago hasta aplicaciones para el control de acceso [46] que permiten el acceso de usuarios con una simple foto capturada por la cámara frontal de un móvil.

No obstante, pese a que, tal y como afirman los autores en [47], la mayoría de los sistemas de reconocimiento facial funcionan correctamente en entornos restringidos, su rendimiento se degrada rápidamente en condiciones no reguladas. Así, el reconocimiento facial en dispositivos móviles es un tema de investigación emergente que sigue cobrando relevancia debido a los nuevos retos que presenta, como variabilidad en las condiciones de captura y limitación de recursos, que dificultan la adaptación las aplicaciones de escritorio convencionales.

Los autores de [6] señalan las dificultades generales derivadas del uso de estas técnicas en dispositivos móviles, que pueden ser causadas por factores extrínsecos o intrínsecos. Los factores extrínsecos están estrechamente relacionados con las condiciones de captura: iluminación, pose, rotación, distancia a la cámara, expresiones u oclusiones. En este sentido, pequeñas variaciones en estos factores pueden alterar la apariencia de la cara y reducir la eficiencia del proceso de

localización. Los factores intrínsecos varían directamente el aspecto de la cara y pueden ser a su vez divididos en cambios intrapersonales e interpersonales. Los cambios intrapersonales incluyen el uso de gafas, bufanda o maquillaje, los cambios en la barba o el bronceado de la piel, y las variaciones morfológicas debidas al cambio de peso o la edad, entre otros. Los cambios interpersonales, como la etnia o el género, podrían afectar al rendimiento del sistema, pero no son determinantes en el tema que nos ocupa.

Por otro lado, un reconocimiento facial robusto requiere un considerable esfuerzo de computación para procesar las imágenes. La mayoría de los dispositivos móviles tienen CPU's que trabajan a menos de 1,5 GHz y no poseen unidad de coma flotante, por lo que las operaciones de coma flotantes son emuladas por la CPU, reduciendo la velocidad de respuesta total. Así mismo, los recursos de memoria en dispositivos móviles también son limitados, por lo que no se recomienda implementar algoritmos que demanden demasiada memoria. Así, muy pocos de los algoritmos generales para el reconocimiento facial son realmente apropiados para el desarrollo de aplicaciones en tiempo real en dispositivos móviles. No obstante, se pueden encontrar algunos trabajos relevantes que han sido publicados recientemente.

Respecto a la etapa de detección de las caras en la imagen, se pueden encontrar dos tipos de técnicas: técnicas de segmentación por el color de la piel [48, 49] y técnicas de extracción de características simples [50, 51]. La mayoría de los trabajos actuales se basan en el algoritmo de Viola-Jones [50], que se ha convertido en el trabajo de referencia desde su publicación en 2001. Se trata de un método englobado en el segundo grupo que utiliza una extracción de características sencilla y es capaz de alcanzar una gran tasa de detección (DR, de sus siglas en inglés) con un procesado extremadamente rápido de imágenes. El detector puede llegar a procesar 15 frames por segundo, lo que le convierte en una opción muy adecuada para aplicaciones en tiempo real. Cabe destacar la aproximación presentada en [52], una optimización software del algoritmo de Viola-Jones para la detección facial en tiempo real en plataformas móviles que, gracias a la reducción de datos y búsqueda y el uso de aritmética fija, permite utilizar exclusivamente el procesador del dispositivo, manteniendo libre el co-procesador hardware.

Respecto a la extracción de las características biométricas de cada cara y su comparación, en la literatura se pueden encontrar diferentes aproximaciones orientadas a resolver los problemas derivados del uso de dispositivos móviles.

Ng et al. [53] presentan un nuevo sistema de verificación para dispositivos móviles que incluye filtros UMACE (Unconstrained Minimum Average Correlation Energy) tolerantes al ruido y la distorsión y transformada FFT (Fixed Point 2D Fast Fourier Transform) en la fase de reconocimiento. Los filtros UMACE mejoraran el rendimiento de los algoritmos de aprendizaje cuando existen variaciones de iluminación, mientras que la aritmética fija reduce el tiempo de computación al 50 %.

Una idea similar se plantea en [54], donde Han et al. proponen un nuevo método multimodal que mejora la detección facial mediante iluminación infrarroja cercana (NIR) y Análisis de Componentes Principales (PCA) basado en enteros para evitar operaciones en coma flotante. El uso de luz infrarroja reduce la sensibilidad de PCA a la luz y las operaciones con aritmética fija reducen a un tercio el tiempo de computación.

Tao y Veldhuis proponen en [55] un método de autenticación que utiliza métricas de subespacios y el clasificador de Parzen combinados con un detector basado en la aproximación de Viola-Jones. El detector es entrenado solo una vez en modo offline con un conjunto de muestras capturadas por exposición exhaustiva del usuario al sensor.

En [56], Faundez-Zanuy et al. abordan el problema de la limitación computacional con una nueva aproximación basada en el uso del dominio transformado Walsh-Hadamard y el clasificador NN (Nearest Neighbour). La transformada WHT (Walsh-Hadamard Transform) puede

Tabla 2.2: Comparativa de técnicas de reconocimiento facial

Referencia	Sensor	Base de datos	Técnica	Resultados obtenidos
[53]	Cámara de un móvil	Privada (720 imágenes)	Filtro UMACE + FFT	EER=8,49 %
[54]	Cámara de infrarrojo cercano	Privada	PCA basado en enteros	EER=14,79 % Tiempo=79,55ms
[55]	Cámara	BioID	Clasificador Parzen	EER=1,2 %
[56]	Cámara	FERET, ORL	WHT + ANN	DCF=5,45
[51]	Cámara	ORL, ETRI	LDA + DCT +ANN	DR=96 % Tiempo=243-412ms
[57]	Cámara	Bayer	GMM	DR=88,5 % Tiempo=52,9ms
[58]	Cámara de alta resolución	Privada (45 imágenes)	Etiquetado reginal + PCA	DR=84,3 %, EER=35 % Tiempo=1,6s
[58]	Cámara de alta resolución	Privada (45 imágenes)	Etiquetado reginal + LDA	DR=94 %, EER=25 % Tiempo=1,6s
[59]	Cámara	O2FN, AR	C-APCDA + CST + ERE	DR=96 %, EER=1,9 %

ser implementada de forma sencilla en aritmética fija y obtiene un buen compromiso entre la demanda de almacenamiento, el tiempo de ejecución y el rendimiento.

Jung et al. proponen en [51] la utilización del brillo relativo entre las distintas partes de la cara, EP (Energy Probability) y LDA (Linear Discriminant Analysis) para extraer las características de la imagen facial, previamente transformada mediante DCT (Discrete Cosine Transform), y un clasificador NN para la comparación de características.

En [48] Hadid et al. presentan un entorno orientado a móviles basado en LBP (Local Binary Pattern) y HI (Histogram Intersection) como medida de disimilitud en la fase de autenticación.

Reng et al. [59] se centran en la mejora de la precisión del alineamiento facial y presentan un nuevo detector de cara y ojos conocido como C-APCDA (Cascade Asymmetric Principal Component Discriminant Analysis). También proponen una nueva aproximación basada en el cálculo de un umbral específico para cada clase a la hora de construir el subsespacio de caras. Las aproximaciones basadas en subespacios ralentizan el tiempo de computación por las multiplicaciones de matrices de gran dimensión que requieren, pero necesitan menos memoria y trabajan mejor cuando las imágenes son de baja resolución. Así mismo, los resultados al utilizar umbrales específicos son mucho mejores que los proporcionados por umbrales globales.

Finalmente, en [58] Dave et al. presentan un análisis de las técnicas de detección y reconocimiento facial más populares implementadas en un teléfono Motorola Droid. La detección facial se lleva a cabo mediante una combinación de segmentación basada en color, procesado morfológico y comparación de patrones, asumiendo condiciones de iluminación correctas, que los usuarios miran a la cámara y que están cercanos a la misma, para simplificar los algoritmos. En casos de mala iluminación o colores de piel oscuros utilizan algoritmos de etiquetado de regiones. Para el reconocimiento facial emplean los esquemas Eigenfaces y Fisherfaces. Para la implementación del esquema Fisherfaces hacen uso del detector facial proporcionado por la API (Application Programming Interface) de Android. Tanto para el esquema de Eigenfaces como el de Fisherfaces el entrenamiento de las matrices KLT y Fisher LDA se realiza de forma separada en un ordenador mediante un programa implementado en MatLab y el resultado del entrenamiento se almacena posteriormente en el dispositivo. Con el objetivo de reducir el tiempo de computación,

las imágenes de alta resolución de este dispositivo son reescaladas, reduciendo su tamaño en un factor de 8, y las matrices de entrenamientoson convertidas a un tipo óptimo para posteriores operaciones.

La Tabla 2.2 muestra una comparativa de todas estas técnicas.

2.4. Biometría multimodal

En los últimos años se ha producido un crecimiento de la tendencia hacia la fusión de información biométrica con el objetivo de mejorar la respuesta de los sistemas tradicionales. La fusión de múltiples modalidades biométricas permite sacar partido de las fortalezas de cada modalidad, compensando algunas de sus limitaciones. En concreto, la explotación de información complementaria incrementa la precisión y reduce la vulnerabilidad, ofreciendo un sistema biométrico más robusto y seguro contra fraude [60].

En general, el término Fusión Biométrica se considera un sinónimo de Biometría Multimodal pero, de acuerdo con [61], incluye dos técnicas generales: Fusión Multimodal, donde la información biométrica se obtiene de diferentes rasgos físicos o de comportamiento, y Fusión Intramodal donde la información biométrica se obtiene del mismo rasgo, pero usando diferentes características, clasificadores o sensores. En el caso particular de la biometría facial, también es posible encontrar información sobre diferentes características usando el mismo sensor.

Atendiendo al módulo del sistema en el que se combina la información biométrica, se puede distinguir entre cuatro niveles de fusión [60]: sensor o datos, función, coincidencia o puntuación y decisión.

La efectividad de un sistema biométrico multimodal generalmente depende del nivel donde se realiza la fusión biométrica: generalmente, cuanto más temprana es la etapa, más efectivo es el sistema, porque la información sobre la biométrica del sujeto a identificar disminuye significativamente con el procesamiento de datos en los diferentes niveles [62]. No obstante, aunque se espera que la fusión a nivel de características proporcione mejores resultados de reconocimiento, la integración a este nivel es difícil de lograr en la práctica porque los conjuntos de características de las diversas modalidades biométricas pueden no ser compatibles. Por otro lado, la fusión en el nivel de decisión se considera rígida debido a la información extremadamente limitada disponible en esta etapa. Por lo tanto, generalmente el nivel preferido por la mayoría de los investigadores para fusionar la información biométrica es el de puntuación, ya que es relativamente fácil obtener y combinar las puntuaciones presentados por las diferentes modalidades

Como resultado de la tendencia hacia la biometría sin contacto y el crecimiento de la fusión biométrica, algunos autores han comenzado a trabajar en la fusión biométrica para dispositivos móviles [63, 64, 65]. Sin embargo, si bien se pueden encontrar trabajos que fusionan la biometría facial con la información proveniente de otros rasgos, apenas hay trabajos relacionados con la fusión intramodal de información facial, y hasta donde llega mi conocimiento, no están orientados a móviles [66, 67, 68].

3

Descripción del sistema biométrico implementado

En este capítulo se explicará en detalle el diseño, la implementación y el funcionamiento del sistema biométrico propueto. Tal y como se comentó en el capítulo de Introducción, el sistema posee la estructura típica de un sistema biométrico de verificación. Por lo tanto, posee cuatro módulos principales: adquisición y preprocesado, extracción de características, comparación y decisión. Además de estos módulos, los sistemas biométricos también constan de una base de datos en la que se almacenan los datos biométricos de los usuarios enrolados en el sistema, junto con un identificador del usuario al que pertenecen. El sistema propuesto posee además un módulo de fusión de información.

3.1. Adquisición y preprocesado

Durante la adquisición, el usuario presenta su muestra biométrica al sistema y los datos brutos son capturados mediante diferentes sensores según la característica a capturar. En el caso de este proyecto, el sensor es la cámara frontal de un dispositivo móvil.

Este módulo lleva a cabo dos tareas principales: la adquisición y el preprocesado de la imagen adquirida. En la adquisición, el sensor captura una imagen de la cara del usuario en un entorno no controlado. A continuación, la imagen será procesada para adaptarla a los requisitos del siguiente módulo: la extracción de características. La fase de preprocesado incluye las tareas de detección de cara en la imagen capturada y recorte de la misma para quedarse únicamente con el rostro del usuario. En caso de haber más de una cara en la imagen, el sistema seleccionará la de mayor tamaño entre todas las detectadas. Además, la imagen recortada será convertida a escala de grises, requisito de los métodos de textura. En la Figura 3.1 se muestra el funcionamiento de este módulo de forma gráfica.

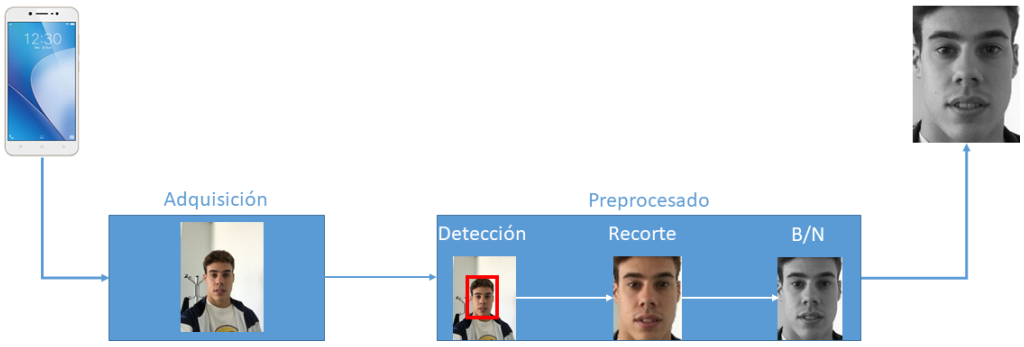

Figura 3.1: Adquisición y preprocesado del sistema biométrico

3.2. Extracción de características

En el módulo de extracción de características, los datos biométricos capturados por el sensor se reducen a un conjunto de características biométricas que son de interés para el reconocimiento del usuario. Estos conjuntos son aproximaciones de los datos recogidos por el sensor y contienen información más discriminatoria e invariable que los datos capturados. A partir de las características extraídas de una o varias muestras se compone el patrón biométrico del usuario, que se utilizará como referencia en futuros accesos. En el proyecto realizado, estos conjuntos de características serán las características geométricas extraídas a partir de los *landmarks* y la descripción de textura obtenida de los patrones binarios locales y los patrones derivativos locales.

A continuación se describirán los diferentes métodos y librerías utilizadas para la extracción de características. En primer lugar se detallará la aproximación basada en *landmarks*. A continuación, se detallarán las aproximaciones holísticas, en las que se analiza la textura de la imagen para extraer información biométrica de la cara. En la Figura 3.2 se ilustra el funcionamiento de este módulo. Tal y como se puede observar en la figura, la salida de la extracción de características serán distancias y ángulos para el método basado en puntos característicos de la cara y descriptores de textura obtenidos a partir de vectores LBP y LDP para los métodos basados en textura.

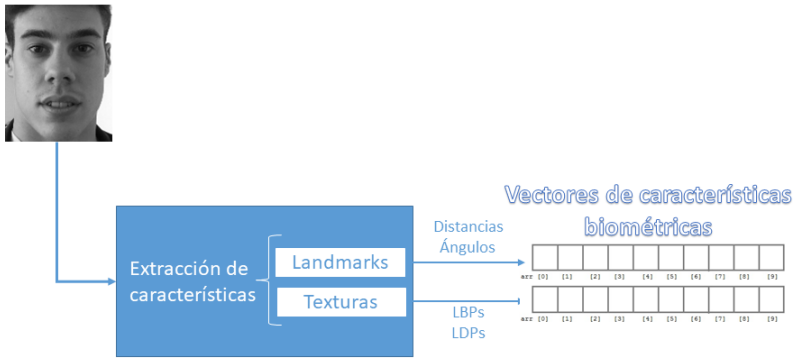

Figura 3.2: Extracción de características del sistema biométrico

3.2.1. Métodos basados en *landmarks*

Los métodos basados en *landmarks* utilizan las relaciónes geométricas entre puntos característicos de la cara. Para ello, es necesario detectar la posición de estos *landmarks* de forma precisa, ya que de esta detección dependen resultados futuros. En este TFM, se han estudiado y comparado dos aproximaciones diferentes: una que utiliza el detector de *landmarks* incluido en la API de Google y una que emplea la librería de libre uso DLIB, en este caso para C++.

En ambos casos, el patrón biométrico de cada usuario incluye distancias entre los puntos detectados. En el caso de la librería DLIB para la detección de *landmarks* se han probado distintas configuraciones de distancias con el fin de obtener la más precisa. Además, para algunas de estas aproximaciones serán analizados los ángulos entre algunos de los puntos.

Detección de *landmarks* mediante la API de Google

Google ofrece una API para implementar aplicaciones de visión artificial en Google basada en Mobile Vision, esta interfaz habilitada tanto para IOS como para Android y ofrece funcionalidades para encontrar objetos en videos y fotografías. Este sistema incluye un paquete específico para detección facial que posee, entre otras, una clase llamada *FaceDetector* y otra llamada *Landmark*. Este paquete puede trabajar con imágenes que incluyan más de una cara. De este modo, al tomar una imagen o video como entrada, puede detectar todas las caras incluidas y devolver los puntos característicos de cada cara. No obstante, tal y como se ha mencionado anteriormente, el módulo de preprocesado seleccionará una única cara, por lo que en nuestro caso simplemente se devolverá un vector de puntos.

En concreto, esta API devuelve los 8 puntos de referencia enumerados a continuación y que se ilustran en la Figura 3.3:

Figura 3.3: Landmarks proporcionados por la API de Google

- Ojo derecho: centro del ojo derecho del usuario.

- Ojo izquierdo: centro del ojo izquierdo del usuario.

- Mejilla derecha: punto medio entre la esquina derecha de la boca y la esquina exterior del ojo derecho. Para caras de perfil completo, este se convierte en el centroide de la base de la nariz, la punta de la nariz, el lóbulo de la oreja derecha y la punta de la oreja derecha.

- Mejilla izquierda: punto medio entre la esquina izquierda de la boca y la esquina exterior del ojo izquierdo. Para caras de perfil completo, este se convierte en el centroide de la base de la nariz, la punta de la nariz, el lóbulo de la oreja izquierda y la punta de la oreja izquierda.

- Base de la nariz: punto medio entre las fosas nasales del usuario.

- Boca derecha: esquina derecha de la boca del usuario donde se juntan los labios.

- Boca izquierda: esquina izquierda de la boca del usuario donde se juntan los labios.

- Boca inferior: centro del labio inferior del usuario.

Como existen 8 puntos, habrá 28 distancias entre ellos. Las 28 distancias extraídas de las imágenes de enrolamiento componen el patrón biométrico de cada usuario.

Detección de *landmarks* mediante la librería DLIB

DLIB es un conjunto de herramientas de matching learning implementado en C++ que fue diseñado para resolver problemas en muchas áreas industriales, incluyendo robótica, dispositivos móviles o visión artificial. DLIB es una librería de código abierto y licencia libre.

DLIB incluye algunas funcionalidades que permiten encontrar rostros humanos en una imagen y estimar su pose. En concreto, la pose devuelve 68 puntos de referencia para cada cara tal y como se muestra en la Figura 3.4. Estos puntos están relacionados con los ojos, las cejas, la nariz, la boca y el contorno de la cara.

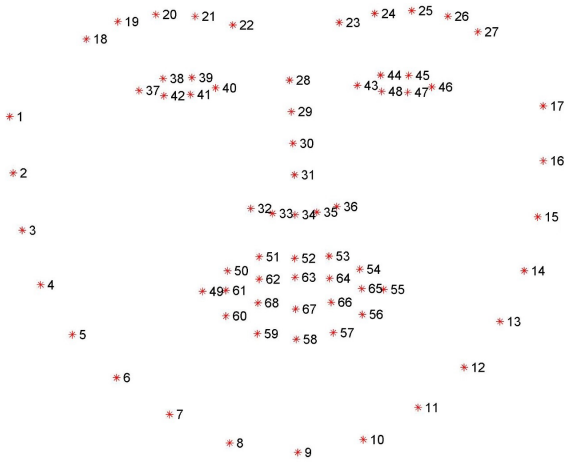

Figura 3.4: Landmarks proporcionados por la librería DLIB

Como hay 68 puntos, existen 2346 distancias entre ellos.

Comparando estos puntos característicos con los puntos de la API de Google, se observa una gran diferencia tanto en el número de puntos como en la precisión de las posiciones devueltas. No obstante, la posición de algunos de los puntos de referencia extraídos por DLIB presenta una alta variabilidad y, por lo tanto, introducen ruido en el patrón biométrico. Por esta razón, se han seleccionado los puntos más representativos y evaluado diferentes combinaciones de distancias.

El análisis visual y los resultados ofrecidos por diferentes pruebas, cuyos resultados se muestran en el Capítulo 4, plantearon que, dependiendo de la postura, los puntos asociados al contorno de la cara no se detectan con precisión y, por tanto, aumentan el error final. Por esta razón, los puntos del contorno de la cara, excepto el mentón, que es un punto esencial, fueron descartados. Respecto a la boca, hay 19 puntos y muchos de ellos son superfluos, porque las distancias entre ellos no proporcionan información relevante. Debido a esto, solo los puntos de referencia numerados como 49, 55, 63 y 67 en la Figura 3.4 se han utilizado para calcular las características incluidas en el patrón biométrico. Analizando los puntos de la nariz, se puede deducir que la información significativa incluida en ellos se concentra en los puntos número 32, 34 y 36 de la Figura 3.4. Finalmente, algunos puntos de los ojos y las cejas fueron eliminados, manteniendo de las cejas los puntos 18, 22, 23 y 27 y de los ojos, se seleccionaron los puntos externos e internos (37, 40, 43 y 46). Además, la posición de la pupila ha sido aproximada como el al punto medio de los puntos de referencia externos e internos de cada ojo. Los puntos seleccionados se muestran en la Figura 3.5.

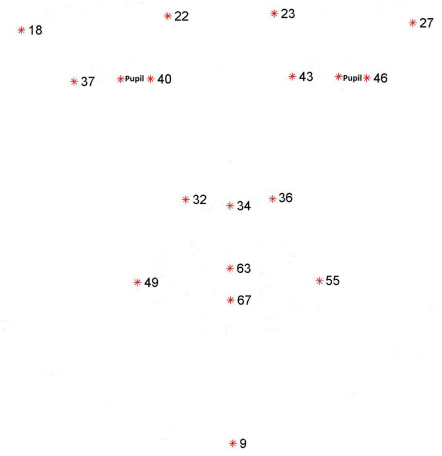

Figura 3.5: Selección de landmarks proporcionados por la librería DLIB

En este caso como hay 18 puntos, existen 171 distancias entre ellos.

Tal y como se verá en los resultados mostrados en el Capítulo 4, se puede deducir que muchas de estas distancias siguen siendo poco relevantes. Por ejemplo, las distancias entre la pupila izquierda y el lado derecho de la boca no son relevantes. Por el contrario, hay algunas distancias que son altamente significativas, proporcionando mucha información sobre el usuario. En consecuencia, se han descartado aquellas distancias poco significativas, manteniendo en el patrón biométrico solamente aquellas más relevantes. La Figura 3.6 muestra la configuración final elegida, en la que la información biométrica del usuario se codifica con 18 de las 171 posibles distancias entre estos 18 puntos.

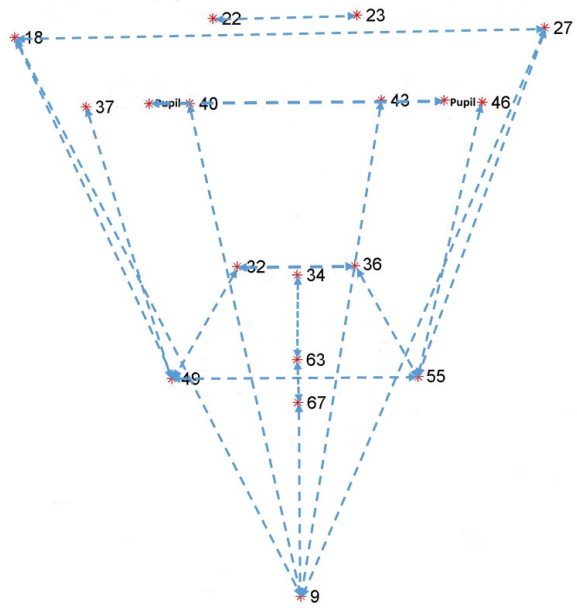

Figura 3.6: Distancias más características de la librería DLIB

Además, para la construcción del patrón geométrico generado a partir de los puntos detectados por DLIB, se han añadido como características biométricas varios ángulos entre algunos de estos puntos. En concreto, se han elegido 18 ángulos característicos de la cara que se pueden ver en las Figuras 3.7, 3.8, 3.8, 3.9, 3.10 y 3.11 que son:

Ángulos desde el mentón (punto 9) hacia:

- las partes exteriores de los ojos (puntos 37 y 46)
- las partes exteriores de la boca (puntos 49 y 55)
- las partes exteriores de la nariz (puntos 32 y 36)
- las partes exteriores de las cejas (puntos 18 y 27)

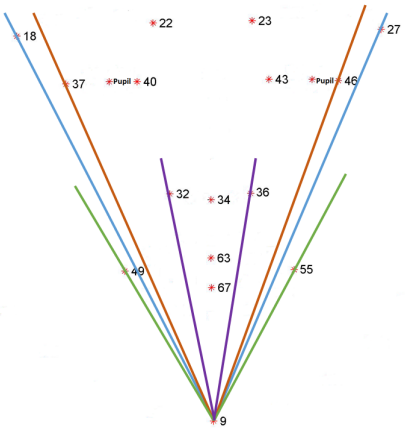

Figura 3.7: Ángulos desde el mentón

tal y como se muestra en la Figura 3.7.

Ángulos desde el exterior derecho de la boca (punto 49) hacia:

- las partes exteriores de las cejas (puntos 18 y 27)
- las partes exteriores de los ojos (puntos 37 y 46)
- el mentón y el exterior derecho de la boca (puntos 9 y 55)

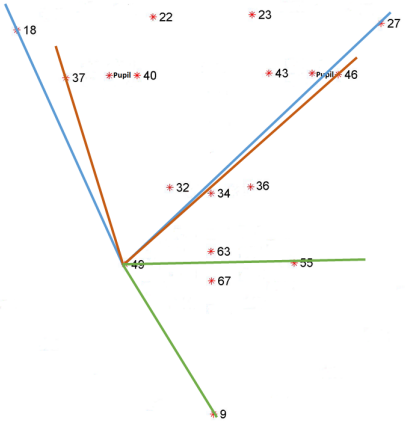

Figura 3.8: Ángulos desde el exterior derecho de la boca

tal y como se muestra en la Figura 3.8.

Ángulos desde el exterior izquierdo de la boca (punto 55) hacia:

- las partes exteriores de las cejas (puntos 18 y 27)
- las partes exteriores de los ojos (puntos 37 y 46)
- el mentón y el exterior izquierdo de la boca (puntos 9 y 49)

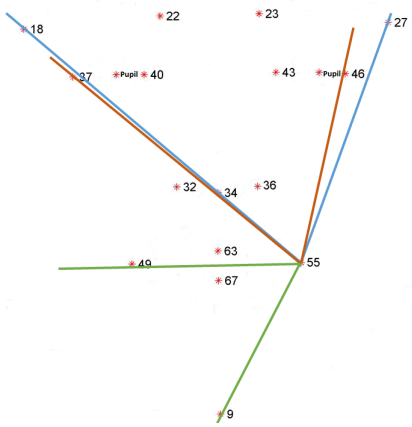

Figura 3.9: Ángulos desde el exterior izquierdo de la boca

tal y como se muestra en la Figura 3.9.

Ángulos desde el centro inferior de la nariz (punto 34) hacia:

- las partes exteriores de las cejas (puntos 18 y 27)
- las partes exteriores de los ojos (puntos 37 y 46)
- las partes interiores de los ojos (puntos 22 y 23)
- las partes exteriores de la boca (puntos 49 y 55)

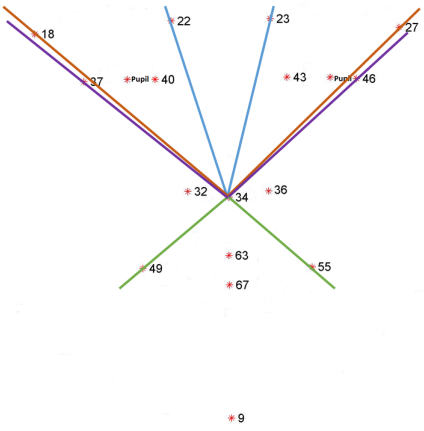

Figura 3.10: Ángulos desde el centro inferior de la nariz

tal y como se muestra en la Figura 3.10.

Ángulos desde el interior del ojo izquierdo (punto 43) hacia:

- las partes exteriores de la nariz (puntos 32 y 36)
- las partes interior y exterior de la ceja izquierda (puntos 23 y 27)

Ángulos desde el interior del ojo derecho (punto 40) hacia:

- las partes exteriores de la nariz (puntos 32 y 36)
- las partes interior y exterior de la ceja derecha (puntos 18 y 22)

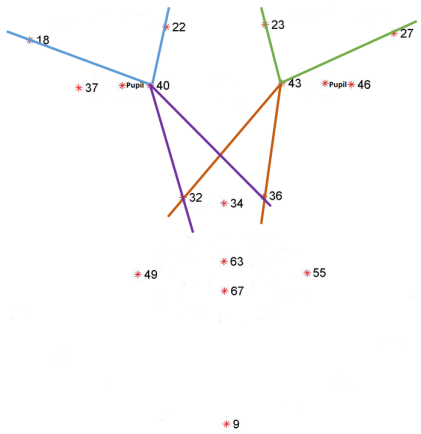

Figura 3.11: Ángulos desde los ojos

tal y como se muestra en la Figura 3.11.

Igualmente, se han añadido algunas distancias características más, obtenidas por experimentación con el objetivo de mejorar la respuesta final del sistema. Las configuraciones que dieron mejores resultados contienen 37 y 47 distancias respectivamente. Los puntos entre los que se calculan cada una de estas distancias están especificados en las Tablas 3.1 y 3.2.

Tabla 3.1: Distancias que componen el patrón geométrico

Número de distancia	**Punto de inicio**	**Punto de fin**
1	18	22
2	18	37
3	18	40
4	18	27
5	18	32
6	18	36
7	18	49
8	23	18
9	23	22
10	23	40
11	23	43
12	23	46
13	23	26
14	23	9
15	37	22
16	37	27
17	37	40
18	37	43
19	37	46
20	37	32
21	37	36
22	37	49
23	37	55
24	43	27
25	43	55
26	32	40
27	32	36
28	32	43
29	32	49
30	32	55
31	32	9
32	49	22
33	49	23
34	49	27
35	49	46
36	49	36
37	49	9

Tabla 3.2: Distancias adicionales incluidas en el patrón geométrico

Número de distancia	Punto de inicio	Punto de fin
38	28	36
39	28	32
40	28	55
41	28	58
42	28	49
43	52	58
44	52	27
45	52	23
46	52	22
47	52	18

3.2.2. Métodos basados en textura

En este apartado se explicará el funcionamiento de los dos métodos holísticos incluidos en el sistema y que utilizan los descriptores de textura LBP (patrones binarios locales) y LDP (patrones derivativos locales).

Patrones binarios locales

Los patrones binarios locales (LBP) son descriptores de textura muy populares presentados por Tim Ojala et al. [69] en el año 1996 cuya robustez, simplicidad y buen rendimiento, tanto en precisión como en eficiencia de cómputo, los hacen adecuados para muchas aplicaciones de visión artificial y procesamiento de imágenes.

En primer lugar, para cada píxel de la imagen se calcula su código LBP, utilizando las intensidades correspondientes a ese píxel y a sus vecinos. En el operador original LBP, se establece un vecindario de 3x3. Para calcular el patrón se evalúa si la intensidad del píxel evaluado es igual, menor o mayor que la de sus vecinos. En caso de ser menor o igual, a ese vecino se le asigna un valor 1, y en caso contrario 0. Después, se concatenan los valores asignados a los vecinos obteniendo como resultado un patrón binario. Este patrón binario, o su equivalente hexadecimal, codifican la información de textura contenida en ese píxel. Este proceso se describe formalmente como:

$$LBP_{code} = \sum_{p=0}^{P-1} S(g_p - g_c)2^p$$

donde g_p corresponde al valor de la intensidad en escala de grises del píxel vecino $(p_0, .., p_{P-1})$, q_c es el valor en escala de grises del píxel central, P es el número de vecinos y S es la función umbral definida como:

$$S(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

La textura de la imagen se representa comúnmente como el histograma de los códigos LBP calculados.

En 2002, Ojala et al. [70] extendieron la definición del operador de LBP a un vecindario de diferentes tamaños. El operador original se deriva a un caso general basado en la simetría circular de una región de P píxeles vecinos dentro de un círculo de radio R. Siguiendo este principio, el operador se denota como $LBP_{P,R}$. De esta forma, el operador de LBP se define por dos

parámetros: P y R, donde P representa el número de píxeles vecinos y controla la cuantificación del espacio angular y R corresponde al radio del círculo y determina la resolución espacial del operador.

En ese trabajo Ojala et al. también observaron que hay algunos patrones binarios que ocurren con más frecuencia en la descripción de la textura: los patrones binarios uniformes. Estos patrones contienen muy pocas transiciones espaciales, no hay más de dos cambios bit a bit 0/1 en el patrón cuando se recorre circularmente. 0000000, 1111111, 00000111 ó 0001100 son ejemplos de patrones binarios uniformes, mientras que 00100110 no es uniforme. En este caso, para calcular el histograma de la imagen, se asigna una etiqueta diferente a cada patrón binario uniforme y se asigna otra al resto de patrones binarios. De esta forma, se obtiene un descriptor de textura más corto sin perder información relevante y, por lo tanto, lo suficientemente representativo de la distribución de las características locales de la imagen. La descripción formal del código es:

$$LBPU_{code} = \begin{cases} \sum_{p=0}^{P-1} S(g_p - g_c)2^p, \ si\ es\ uniforme \\ P{+}1, \ de\ otra\ manera \end{cases}$$

donde g_p corresponde al valor de la intensidad en escala de grises del píxel vecino $(p_0, .., p_{P-1})$, q_c es el valor en escala de grises del píxel central, P es el número de vecinos y S es la función umbral definida como:

$$S(x) = \begin{cases} 1, \ x{\geq}0 \\ 0, \ x{<}0 \end{cases}$$

Aunque los histogramas de textura podrían usarse directamente como el vector de características correspondiente al usuario del sistema biométrico, con el objetivo de agregar alguna información global, la imagen se divide en varias regiones y se calcula un histograma para cada sub-imagen. Finalmente, los histogramas se concatenan para componer el vector de características biométricas final que describe la cara.

En este TFM se han probado diferentes valores para el número de vecinos, el radio y el tamaño de las subregiones en las que se divide la imagen. En concreto, respecto al número de vecinos se han probado configuraciones de 8 o 16 vecinos, para el radio se han tomado valores 1, 2 y 3 y los tamaños de las subregiones testeadas han sido: 8x8, 16x16, 32x32 y 64x64.

Patrones derivativos locales

Los Patrones Derivativos Locales (LDP) son descriptores de textura de alto orden originalmente propuestos para el reconocimiento facial con el objetivo de capturar información discriminativa más detallada que los patrones locales de primer orden como los utilizados en LBP. Un LDP codifica características mediante patrones direccionales basados en las variaciones derivadas locales de orden (n-1) obtenidas mediante una función de codificación binaria.

Dada una imagen I(Z), siendo Z_0 un píxel en I(Z), y Z_i, i = 1,...,8 siendo los píxeles vecinos alrededor de Z_0 donde Z_1 es el vecino superior a la izquierda, Z_2 es el vecino superior, etc. El orden direccional enésimo en la dirección α en $Z = Z_0$ se define por la ecuación:

$$LDP_{\alpha}^{n}(Z_0) = \{f(I_{\alpha}^{n-1}(Z_0), I_{\alpha}^{n-1}(Z_1), I_{\alpha}^{n-1}(Z_2), ..., , I_{\alpha}^{n-1}(Z_0), , I_{\alpha}^{n-1}(Z_s)\}$$

donde f(...) es una función de codificación binaria que determina los tipos de transiciones entre patrones locales e $I^{n-1}(Zj)$ es el la derivada de orden en la dirección α en $Z = Z_j$ siendo j=0,...,8 y $\alpha = 0, 45, 90\ y\ 135$ grados.

Concretamente, $f(I^{n-1}(Z_0), I^{n-1}(Z_j))$ codifica como patrones binarios las transiciones entre gradientes de orden $(n-1)$ y se define de la siguiente manera:

$$f(I_{\alpha}^{n-1}(Z_0), I_{\alpha}^{n-1}(Z_i)) = \begin{cases} 0, \ si\ I_{\alpha}^{n-1}(Z_i) \cdot I_{\alpha}^{n-1}(Z_0) > 0 \\ 1, \ si\ I_{\alpha}^{n-1}(Z_i) \cdot I_{\alpha}^{n-1}(Z_0) \leq 0 \end{cases}, \ i = 1, ..., 8$$

El LDP de orden enésimo es una concatenación de LDPs direccionales en cuatro direcciones con una resolución de 45 grados definida de acuerdo con la ecuación:

$$LDP^n(Z) = LDP^n_\alpha(Z) \mid \alpha = 0, 45, 90 \; y \; 135$$

Una vez calculados los códigos binarios, el algoritmo funciona de la misma manera que el enfoque basado en LBP y la textura de la imagen se representa mediante el histograma de los códigos LDP. Del mismo modo, con el objetivo de agregar alguna información global, la imagen se divide en varias regiones se calcula el histograma para cada sub-imagen y se concatenan los histogramas.

3.3. Comparación

En el módulo de comparación se confrontan dos muestras biométricas con el objetivo de detectar el grado de similitud entre ellas. Para ello, se ha utilizado un clasificador basado en distancias. En concreto para la aproximación geométrica se han utilizado las siguientes distancias: Euclidea, Manhattan y Chebyshev. En el caso de las aproximaciones holísticas se utilizará la distancia Chi-Square. Por lo tanto, la salida de este módulo será una puntuación que refleje las diferencias entre la nueva muestra y el patrón contra el que se compara. En la Figura 3.12 se puede observar el módulo de comparación del sistema.

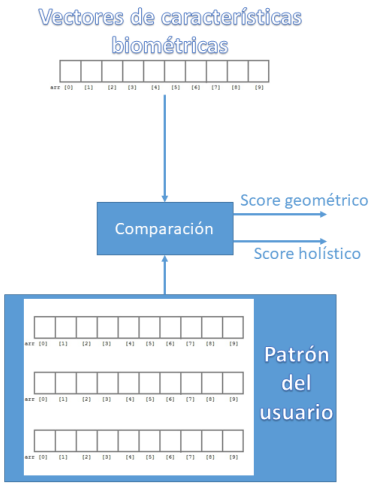

Figura 3.12: Comparación del sistema biométrico

- Distancia Euclídea: es la distancia ordinaria entre dos vectores de un espacio euclídeo. La definición general de la distancia euclídea entre los vectores $P = (p_1, p_2, ..., p_n)$ y $Q = (q_1, q_2, ..., q_n)$ del espacio euclídeo n-dimensional se define como :

$$d_{Euclidea}(P, Q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + ... + (p_n - q_1 n)^2} \; = \; \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2}$$

- Distancia Manhattan: es la distancia entre dos vectores del espacio euclídeo en una cuadrícula basada en una ruta estrictamente horizontal y/o vertical (es decir a lo largo de las líneas de la cuadrícula). Es la suma simple de las componentes horizontales y verticales.

$$d_{Manhattan}(X_1, X_2) = ||X_1 - X_2|| = \sum_{i=1}^{n} |X_{1i} - X_{2i}|$$

- Distancia Chebyshev: es una medida definida en un espacio vectorial donde la distancia entre dos vectores es la mayor de sus diferencias a lo largo de cualquier dimensión de coordenadas. También se conoce como distancia de tablero de ajedrez, ya que en el juego de ajedrez el número de movimientos necesarios para pasar de un cuadrado a otro es igual a la distancia Chebyshev entre los centros de los cuadrados.

$$d_{Chebyshev}(X_1, X_2) = max(|X_{1i} - X_{2i}|)$$

- La distancia Chi-Square viene definida por la siguiente fórmula:

$$d_{Chi-Square}(X_1, X_2) = \sum_{i=1}^{n} \frac{(X_{1i}-X_{2i})^2}{(X_{1i}+X_{2i})}$$

3.4. Fusión

Tras el módulo de comparación, con el objetivo de evaluar si la combinación de información geométrica y holística permite mejorar la precisión del sistema, en alguno de los experimentos se realizará la fusión entre la información ofrecida por ambos métodos (Figura 3.13).

Tal y como se ha comentado en el Estado del Arte, el enfoque más comúnmente utilizado para la fusión biométrica, la fusión a nivel de puntuación debido a su buen rendimiento, simplicidad y al fácil acceso y combinación de las puntuaciones proporcionadas por las diferentes modalidades biométricas. En este nivel, las puntuaciones proporcionadas por las diferentes técnicas biométricas se combinan para obtener una puntuación única que se utilizará para tomar la decisión final sobre la identidad del usuario.

En este TFM se realizará la fusión intramodal a nivel de puntuación o *scores*. Los resultados obtenidos por el módulo de comparación al evaluar las características geométricas y holísticas serán fusionadas en un único *score* enviado al módulo de decisión. Ya que las distribuciones de *scores* de las distintas técnicas están en espacios de características distintos, es necesario realizar previamente una normalización que transforme todos los *scores* a un dominio común. Para ello se ha decidido utilizar la técnica de normalización min-max que permite mantener la distribución original salvo por un factor de escalado, transformando dichos *scores* a un dominio $[0, 1]$.

Así, dado el conjunto de *scores* de una modalidad biométrica (S_i), con $i = 1, ..., N$ y N el número de *scores*, los *scores* normalizados vienen dados por la ecuación:

$$\widetilde{S}_l = \frac{S_i - min(\{S_i\}_{i=1}^{N})}{max(\{S_i\}_{i=1}^{N}) - min(\{S_i\}_{i=1}^{N})}$$

Una vez normalizados los *scores* de las distintas modalidades biométricas, son fusionados para obtener un único *score* que permita tomar una decisión final sobre la identidad del usuario. En términos matemáticos, esto significa que las puntuaciones s_i, con $i = 1, ..., N$ y N el número de modalidades biométricas, se mapean desde $R^N \rightarrow R$ mediante una función f para obtener una puntuación única S:

$$S = f(w_1 s_1, ..., w_n s_n)$$

donde los factores w_i ponderan la influencia de cada *score* en la fusión. Los pesos son optimizados mediante un algoritmo genético. Este algoritmo también permite probar diferentes reglas para la fusión y decidir cual tiene un mejor rendimiento. En concreto, se han probado cuatro reglas ponderadas: mínimo, máximo, suma y producto.

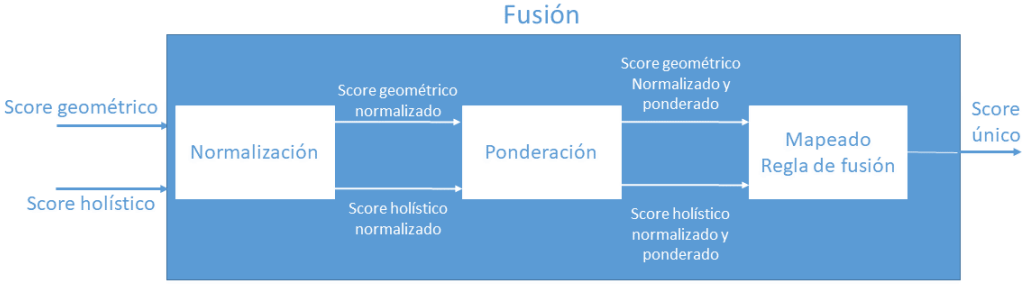

Figura 3.13: Fusión del sistema biométrico

3.5. Decisión

Finalmente, el módulo de decisión, tal y como su nombre indica, es el encargado de tomar una decisión sobre la identidad del usuario. En este módulo, la puntuación resultante del módulo anterior se compara con un umbral previamente fijado y dependiendo de si es mayor o menor que ese umbral, esa muestra será considerada como perteneciente a ese usuario o no. De esta manera, la salida de este módulo, y por lo tanto del sistema biométrico, será aceptar o rechazar la identidad reclamada por el usuario. La Figura 3.14 muestra este proceso de forma gráfica.

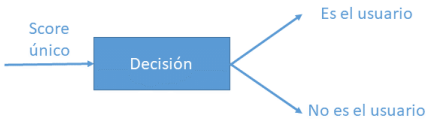

Figura 3.14: Decisión del sistema biométrico

4

Experimentos Realizados y Resultados

En este capítulo se presentarán los experimentos realizados y los resultados que se han obtenido con ellos. Además se detallarán las bases de datos utilizadas y el protocolo de evaluacón biométrico seguido para la evaluación del sistema.

4.1. Bases de datos

Para la evaluación del sistema propuesto en este TFM se han empleado diversas bases de datos que presentan diferentes características tanto en número de usuarios e imágenes por usuario como en la variabilidad de su apariencia o las condiciones ambientales. A continuación se describen los distintos conjuntos de imágenes empleados.

4.1.1. Face-Scrub

El conjunto de datos FaceScrub [71] se creó a partir de imágenes faciales detectadas automáticamente en imágenes obtenidas mediante búsquedas de figuras públicas en Internet. Los resultados se verificaron y limpiaron manualmente para asegurar que cada subconjunto de imágenes pertenece a la celebridad buscada. La base de datos comprende un total de 106.863 imágenes faciales de 530 celebridades masculinas y femeninas, con alrededor de 200 imágenes por persona. Tal cantidad de imágenes la hace una de las bases de datos públicas más grandes disponibles por el momento. Las imágenes no presentan ningún tipo de restricción, por lo que se puede observar una gran variabilidad de poses, iluminación y expresiones. Además, al tratarse de una base de datos de figuras públicas, también son reseñables los cambios en el aspecto físico de los usuarios (pelo, barba, gafas,...) Algunas imágenes de muestra se pueden ver en la Figura 4.1.

Figura 4.1: Imágenes de muestra de la base de datos Face-Scrub

4.1.2. Base de datos propietaria del grupo gb2s

Esta base de datos contiene vídeos de 60 personas capturadas usando un dispositivo móvil en un ambiente interior sin restricciones de fondo, pose o iluminación. Las imágenes también presentan cierta variación en las expresiones de los usuarios. Se han extraído 250 imágenes por cada usuario que corresponden con los *frames* que componen cada uno de los vídeos. Algunas imágenes de muestra se pueden ver en la Figura 4.2.

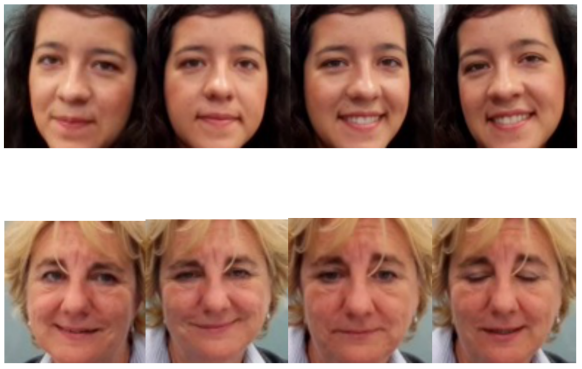

Figura 4.2: Imágenes de muestra de la base de datos del grupo GB2S

4.1.3. Base de datos facial BioID

La base datos BioID [72] es una base de datos pública y de libre descarga. Esta base de datos ha sido creada y publicada para ayudar a investigadores que trabajan en el área de la detección y el reconocimiento facial, permitiendo comparar la calidad de sus algoritmos con otros. Durante la captura se hizo especial hincapié en mantener condiciones del mundo real. Por lo tanto, el conjunto de pruebas presenta una gran variedad de iluminación, fondo, tamaño de cara, pose y expresiones. Esta base de datos incluye 1521 imágenes en escala de grises correspondiente a 23 personas diferentes, de tamaño 384x286 pixeles. Algunas imágenes de muestra se pueden ver en la Figura 4.3.

Figura 4.3: Imágenes de muestra de la base de datos BioID

4.2. Protocolo de Evaluación

En este proyecto se ha utilizado un protocolo basado en las definiciones sugeridas en la norma ISO/IDE 19795, que presenta los requisitos y mejores prácticas científicas para realizar pruebas de rendimiento técnico, así como las directrices proporcionadas en [3] para la evaluación del sistema biométrico desarrollado en este TFM.

4.2.1. Definiciones de la norma ISO/IDE 19795

A continuación se detallan algunas definiciones importantes incluidas en la norma ISO/IDE 19795 con respecto a aplicaciones biométricas, datos, interacción del sistema y evaluación:

- Muestra: medidas biométricas del usuario como resultado del sistema de captura de datos.
- Características: representación digital de la información biométrica extraída a partir de una muestra.
- Modelo: medida de referencia del usuario almacenada en el sistema basada en características extraídas de sus muestras.
- Puntuación de comparación o semejanza: medida de similitud entre las características derivadas de una muestra y un modelo almacenado, medida de lo cerca que están estas características al modelo de referencia de un usuario.
- Presentación: presentación de una única muestra biométrica por parte de un usuario.
- Intento: captura de una (o una secuencia de) muestras biométricas por el sistema.
- Transacción: secuencia de intentos por parte de un usuario para realizar operaciones con el sistema. Hay tres tipos de transacciones: secuencia de enrolamiento, que tiene como respuesta un enrolamiento o un fallo de enrolamiento; secuencia de verificación, que da como resultado una decisión de verificación; o secuencia de identificación, lo que resulta en una decisión de identificación.
- Decisión de verificación: determinación de la validez de la identidad del usuario en el sistema.
- Intento genuino: intento de buena voluntad de un usuario para hacer coincidir su nueva muestra con su modelo previo almacenado.

- Intento impostor zero-effort: intento en el que un individuo presenta sus propias características biométricas como si estuviera intentando una verificación exitosa con su propio patrón, pero la comparación se realiza contra el patrón de otro usuario.

- Evaluación de tecnología: evaluación *offline* de uno o más algoritmos para la misma modalidad biométrica utilizando un abanico de muestras preexistentes o especialmente recogidas para ese fin, idealmente por un sensor universal. Los resultados obtenidos utilizando estas muestras dependerán tanto de las condiciones ambientales como de la población. Como las muestras son fijas, los resultados de las pruebas de evaluación de la tecnología son repetibles.

- Evaluación *offline*: ejecución del enrolamiento y el acceso, por separado del envío de la muestra. La recopilación de una base de datos de imágenes o señales para el enrolamiento *offline* y el cálculo de la puntuación de comparación permite un mayor control sobre parámetros como los intentos de acceso o el número de imágenes utilizadas para comparar el patrón biométrico del usuario. Las pruebas tecnológicas siempre implican el almacenamiento de datos para un posterior procesamiento *offline*.

- Índice de fallo de enrolamiento (FTE): porcentaje de que alguien no sea registrado a causa de un fallo a la hora de crear un patrón. El FTE observado se mide sobre la población de enrolamiento del conjunto de prueba. El FTE previsto/esperado se aplicará a toda la población objetivo.

- Índice de fallo de adquisición (FTA): porcentaje de intentos de verificación o identificación para los cuales el sistema no logra capturar o ubicar una imagen o señal de calidad suficiente.

- Tasa de no falsa coincidencia (FNMR): probabilidad de que dos muestras pertenecientes a la misma persona sean clasificadas como no pertenecientes a la misma clase.

- Tasa de falsa coincidencia (FMR): probabilidad de que dos muestras pertenecientes a dos personas distintas sean clasificadas como pertenecientes a la misma clase.

- Tasa de falso rechazo (FRR): probabilidad de que un usuario que está autorizado sea rechazado a la hora de intentar acceder al sistema.

- Tasa de falsa aceptación (FAR): probabilidad de que un usuario no autorizado sea aceptado por el sistema.

4.2.2. Protocolo de Evaluación de la tecnología

Al diseñar el protocolo de evaluación de tecnología para el sistema de verificación biométrica presentado anteriormente, se han tenido en cuenta las siguientes consideraciones:

- Bases de datos: se usarán diferentes bases de datos para probar los algoritmos. Sus características respecto a las condiciones de captura, la cantidad de usuarios, las sesiones o el número de imágenes por usuario y sesión pueden variar, por lo que el protocolo de evaluación debe ser lo suficientemente flexible. Se realizará una evaluación por separado para cada base de datos.

- Enrolamiento: para ser reconocido por el sistema, un usuario debe estar previamente registrado. Con este fin, un subconjunto de las imágenes de cada usuario se utiliza para obtener el patrón biométrico requerido para inscribir al usuario en el sistema.

- Acceso al sistema: durante la etapa de acceso al sistema se diferencia entre usuarios genuinos e impostores con el objetivo de evaluar la respuesta del sistema en diferentes escenarios. De esta forma, cuando un usuario es considerado genuino, los otros usuarios se consideran impostores, siguiendo un esquema zero-effort.

El protocolo de evaluación seguido se compone de tres partes. En primer lugar, es necesario separar los conjuntos de datos originales para permitir tareas de enrolamiento y acceso. A continuación, se calculan las diferencias entre las características biométricas extraídas de las muestras de enrolamiento y las provenientes de las muestras de acceso. Finalmente, se calculan las métricas del rendimiento del sistema.

1. Organización del conjunto de datos: Para cada conjunto de datos, las imágenes se dividen en subconjuntos de Entrenamiento y Acceso.

2. Cálculo de puntuaciones: Una vez que se han definido los conjuntos de datos de entrenamiento y acceso, estos últimos se dividen en muestras genuinas e impostoras correspondientes a usuarios auténticos e impostores, respectivamente. Luego, se ejecuta la siguiente secuencia de acciones para evaluar el sistema:

1. El modelo o patrón biométrico de cada usuario se crea a través de sus muestras de enrolamiento.
2. Se comparan las muestras de acceso de cada usuario con su patrón biométrico, proporcionando una lista de puntuaciones genuinas para estos intentos.
3. Las muestras de usuarios impostores se comparan con el modelo biométrico de los usuarios que están intentando suplantar, proporcionando una lista de puntuaciones impostoras.
4. Ambas puntuaciones se utilizan para obtener ciertas métricas, que brindan información sobre el rendimiento del sistema.

3. Evaluación: Se calcularán las siguientes métricas:

- Tasa de fallo de enrolamiento (FTE): Proporción de la población para la cual el sistema no completa el proceso de enrolamiento.
- Tasa de fallo de adquisición (FTA): Proporción de intentos de verificación o identificación para los cuales el sistema no logra capturar o ubicar una imagen o señal de calidad aceptable.
- Tasa de falsa coincidencia (FNMR): Proporción de intentos genuinos falsamente declarados como no coincidentes con el modelo del usuario que suministra la muestra.
- Tasa de falsa coincidencia (FMR): Proporción de muestras de intentos de impostor zero-effort falsamente declaradas como coincidentes con el modelo del usuario que está siendo suplantado.
- Tasa de falso rechazo (FRR): proporción de operaciones de verificación genuinas que se negarán de forma incorrecta.

  FRR = FTA + FNMR x (1 - FTA)
- Tasa de falsa aceptación (FAR): proporción de operaciones impostoras que se aceptarán incorrectamente.

  FAR = FMR x (1 - FTA).

La FTE se puede considerar como una medida de calidad sobre como de bueno es el rendimiento de un algoritmo de procesamiento para cierto tipo de imágenes. La tasa de falsa aceptación y la tasa de falso rechazo dependen del umbral de aceptación del sistema biométrico, que se fija de acuerdo con una política de seguridad. Una política muy común en los sistemas biométricos es ubicar el umbral de aceptación en el valor donde FAR y FRR son iguales. Este valor generalmente se denomina Tasa de Error Igual (EER) y es una métrica comúnmente aceptada para cuantificar y comparar el rendimiento de un algoritmo biométrico. En consecuencia, los resultados de evaluación del sistema propuesto se proporcionarán en términos de EER.

4.3. Experimentos y resultados

Una vez estudiados y comprendidos los métodos y las librerías descritos en el capítulo 3, se han realizado numerosos experimentos para evaluar su rendimiento analizando diferentes parámetros. Para esta evaluación se han usado las 3 bases de datos explicadas en este capítulo. Dadas las características de cada base de datos, tales como número de usuarios, número de imágenes por usuario, condiciones externas o tamaño de las imágenes, los resultados ofrecidos por cada una de ellas son difícilmente comparables entre sí, por lo que se presentarán y analizarán por separado.

4.3.1. Aproximación geométrica

En este apartado se mostrarán los resultados obtenidos empleando el método de extracción de características geométricas y la distancia Euclidea como comparador. Los resultados obtenidos utilizando las distancias Manhattan y Chebyshev se podrán consultar en el Anexo A.

Experimento Face-Scrub

En primer lugar se ha realizado un análisis visual de los resultados ofrecidos por el sistema en la fase de detección de puntos característicos. En la figura 4.4 se muestran los resultados obtenidos por las funciones de la API de Google en algunas imágenes de ejemplo. En la Figura 4.5 se muestran los resultados ofrecidos opr la librería DLIB para esas mismas imágenes.

De acuerdo con el protocolo de evaluación biométrica detallado anteriormente se ha dividido el conjunto de imágenes de cada usuario en:

- 70 % para entrenamiento y 30 % para testear

Figura 4.4: Landmarks, API Google , Face-Scrub

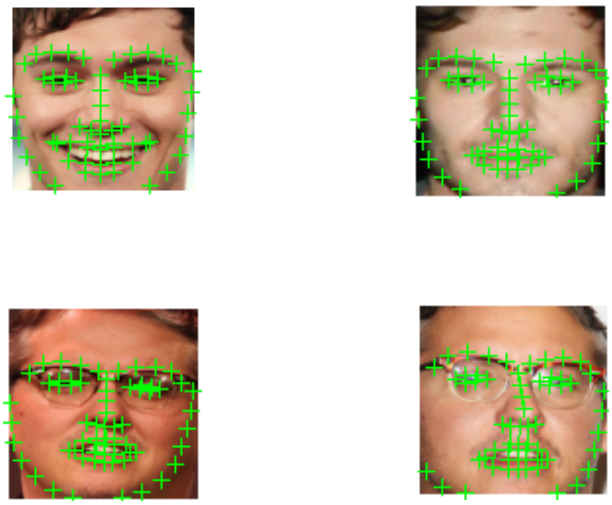

Figura 4.5: Landmarks, DLIB, Face-Scrub

El número de imágenes no es igual en cada usuario por lo que los conjuntos de entrenamiento y de test tendrán diferente número de imágenes dependiendo del usuario. Como se ha comentado en el apartado 3.2.1, se han probado diferentes configuraciones para observar el rendimiento e intentar mejorar las tasas de los algoritmos. En la Figura 4.6 se muestran los 18 puntos más característicos de la cara para la librería DLIB. Se observa que no existen puntos que estén incorrectamente colocados en la cara.

Figura 4.6: 18 Landmarks, DLIB, Face-Scrub

Los resultados obtenidos en la evaluación biométrica para estos dos métodos se muestran en el Tabla 4.1.

Se puede observar que los resultados no son nada satisfactorios en términos de EER. Los algoritmos no parece que sean lo suficientemente buenos para poder verificar al usuario correctamente pues existe entre un 63 % y un 65 % de probabilidad de verificar al usuario correctamente ya que , cuando se utilizan todos los puntos devueltos por los extractores de características y todas las posibles distancias entre ellos, se obtiene un error de más del 37 % y 35 % para el caso de la API de Google y DLIB respectivamente. Además, si nos centramos en la API de Google, al tener tan pocos puntos es muy difícil mejorar las prestaciones de este sistema analizando posibles combinaciones. No obstante, con la librería DLIB es posible hacer diferentes pruebas seleccionando los puntos más representativos y las distancias más características para intentar conseguir mejores prestaciones tal y como se ha explicado en el Capítulo 3.

Al realizar esta selección de puntos se puede observar que el EER se reduce más de un 10 %. Tal y como se esperaba, no se necesitan muchos puntos característicos si no aquellos que realmente diferencian a las personas. La selección de los mejores 18 puntos y las 18 distancias más representativas, hace que 3 de cada 4 individuos son verificados correctamente. Sigue sin ser una tasa realmente buena, pero teniendo en cuenta la resolución de las imágenes de esta base de datos tras la etapa de preprocesado (69x69) es una tasa más que aceptable.

Tabla 4.1: Resultados para la base de datos Face-Scrub con API Google y DLIB

Método de extracción de características	**%imágenes entrenamiento**	**%imágenes test**	**EER (Euclidea)**
API Google 8 puntos 28 distancias	70	30	37,55 %
DLIB 68 puntos 2346 distancias	70	30	36,73 %
DLIB 18 puntos 18 distancias	70	30	25,64 %

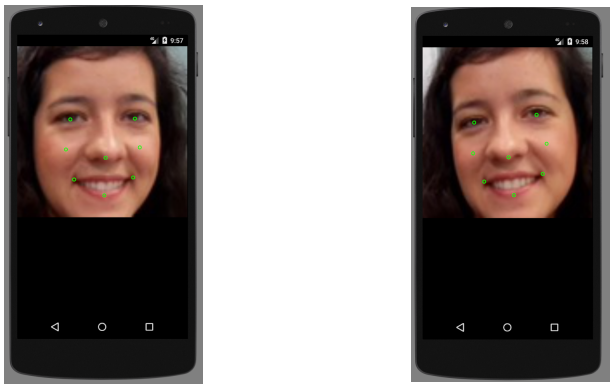

Figura 4.7: Landmarks, API Google , Base de datos GB2S

Figura 4.8: Landmarks, DLIB, Base de datos GB2S

Experimento base de datos propiedad del grupo gb2s

Al suponer el tamaño reducido de las imágenes como una posible causa de los malos resultados, en los experimentos realizados con esta base de datos se han empleado diferentes resoluciones de imagen: 128x128, 500x500 y 890x890.

De forma análoga al experimento anterior, se ha realizado un análisis visual de los resultados proporcionados por los detectores de puntos característicos. Las conclusiones observadas en cuanto a precisión de los mismos se mantienen para esta base de datos. En las Figuras 4.7 y 4.8 se muestran algunas imágenes de la base de datos con los puntos obtenidos.

Los resultados para esta base de datos con imágenes de 128x128 píxeles y mismas configuraciones de puntos y distancias que en el anterior experimento se muestran en la Tabla 4.2. Además, se han incluido pruebas que utilizan también los ángulos descritos en la Sección 3.2.1 del Capítulo 3 como características faciales.

Tabla 4.2: Resultados para la base de datos del grupo gb2s con imágenes de tamaño 128x128

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Euclidea)**
API Google 8 puntos 28 distancias	70	30	33,45 %
DLIB 68 puntos 2346 distancias	70	30	30,74 %
DLIB 18 puntos 37 distancias	70	30	20,86 %
DLIB 18 puntos 37 distancias 18 ángulos	70	30	19,87 %
DLIB 18 puntos 47 distancias 18 ángulos	70	30	19,87 %

Tabla 4.3: Resultados para la base de datos del grupo gb2s con imágenes de tamaño 500x500

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Euclidea)**
API Google 8 puntos 28 distancias	70	30	12,09 %
DLIB 68 puntos 2346 distancias	70	30	7,20 %
DLIB 18 puntos 37 distancias	70	30	5,89 %
DLIB 18 puntos 37 distancias 18 ángulos	70	30	5,56 %
DLIB 18 puntos 47 distancias 18 ángulos	70	30	5,19 %

Tabla 4.4: Resultados para la base de datos del grupo gb2s con imágenes de tamaño 890x890

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Euclidea)**
API Google 8 puntos 28 distancias	70	30	11,19 %
DLIB 68 puntos 2346 distancias	70	30	7,15 %
DLIB 18 puntos 37 distancias	70	30	5,70 %
DLIB 18 puntos 37 distancias 18 ángulos	70	30	5,62 %
DLIB 18 puntos 47 distancias 18 ángulos	70	30	5,16 %

Se puede observar una clara mejoría de estos resultados respecto a los obtenidos con la anterior base de datos. En este caso, se ha llegado a obtener una tasa de verificación correcta por encima del 80 %. Es decir, para esta base de datos, 1 de cada 5 se verificaría incorrectamente. Aun así, el tamaño de las imágenes puede estar afectando a la variabilidad entre los diferentes usuarios. Por lo que se hicieron pruebas adicionales con imágenes de tamaño: 500x500 y 890x890. Los resultados para ambos tamaños se muestran en las Tablas 4.3 y 4.4.

En ambos casos se observa claramente la mejora de resultados tanto utilizando, la API de Google como la librería DLIB, para extraer los puntos faciales. En el caso de la API de Google a pesar de no tener demasiados puntos ni demasiadas distancias se puede observar que el mejor porcentaje de verificación se sitúa en torno al 88 %, mientras que en el caso de DLIB el porcentaje de verificación correcta ronda el 95 %. Es evidente que el tamaño de la imagen influye claramente, aunque se observa que la diferencia entre 500x500 y 890x890 tampoco es demasiado significativa.

Experimento base de datos BioID

De forma análoga al experimento anterior, se ha realizado un análisis visual de los resultados proporcionados por los detectores de puntos característicos. Las conclusiones observadas en cuanto a precisión de los mismos se mantienen para esta base de datos.En las Figuras 4.9 y 4.10 muestran ejemplos de imágenes de esta base de datos con los *landmarks* de la API de Google y de la librería DLIB.

Tras analizar los resultados obtenidos por anteriores bases de datos, en este caso se han replicado únicamente las pruebas que incluyen la API de Google y aquellas que ofrecen los mejores resultados para DLIB. Además, dado el tamaño original de las imágenes (384x286 píxeles) se ha empleado solamente el tamaño de imagen 128x128. En la Tabla 4.5 se muestran los resultados.

Los resultados ofrecidos por esta base de datos confirman las afirmaciones realizadas en los experimentos anteriores. La librería DLIB devuelve mejores resultados que la API de Google, en todas las configuraciones superior al 10 % de mejora. Además, se observa que la mejor configuración de DLIB coincide con la mejor configuración del experimento anterior. Finalmente, decir que para ser un tamaño relativamente pequeño de imagen, se obtienen muy buenos resultados, aproximándose al 10 % de EER.

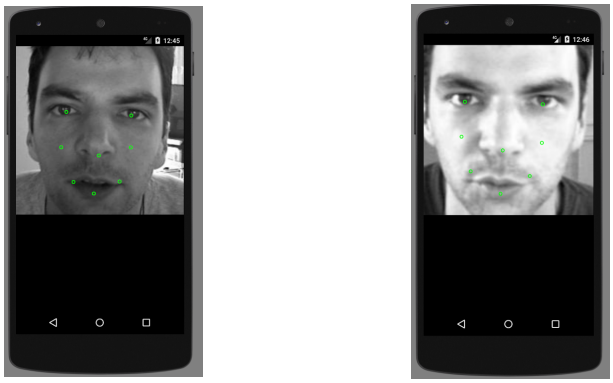

Figura 4.9: Landmarks, API Google, Base de datos BioID

Figura 4.10: Landmarks, DLIB , Base de datos BioID

Tabla 4.5: Resultados para la base de datos BioID con imágenes de tamaño 128x128

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Euclidea)**
API Google 8 puntos 28 distancias	70	30	25,06 %
DLIB 68 puntos 2346 distancias	70	30	14,06 %
DLIB 18 puntos 37 distancias	70	30	14,93 %
DLIB 18 puntos 37 distancias 18 ángulos	70	30	12,21 %
DLIB 18 puntos 47 distancias 18 ángulos	70	30	11,87 %

4.3.2. Aproximación holística

A continuación se detallarán los resultados obtenidos por los métodos basados en texturas faciales. Como se ha explicado en el capítulo 3 el sistema incluye 2 métodos diferentes para describir la textura de las caras:

- Patrones binarios locales (LBP)

- Patrones derivativos locales (LDP)

Para evaluar el rendimiento de estas aproximaciones y obtener resultados comparables se han utilizado las mismas bases de datos que en el experimento anterior.

Experimento Face-Scrub

Por consecuencia con al evaluación geométrica, se ha mantenido la misma distribución de imágenes para los conjuntos de entrenamiento y test: 70 % y 30 % respectivamente. Los resultados obtenidos por el descriptor LBP se muestran en la Tabla 4.6.

Al igual que en la aproximación geométrica, los resultados son bastante desfavorables, ya que las mejores tasas son del 76 %. Se sospecha, que tanto el tamaño de las imágenes (69x69) como la variabilidad de las imágenes pueden ser la causa de los malos resultados. A la vista de los resultados se decidió no probar el descriptor LDP, ya que esta aproximación parece bastante sensible a cambios bruscos de pose, expresión y apariencia como los presentados en esta base de datos.

Tabla 4.6: Resultados LBP Face-Scrub, Subregiones de 8x8, radio 1 y 8 vecinos

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Chi-Square)**
LBP Subregiones=8x8 Radio=1 Vecinos=8	70	10	23,98 %

Tabla 4.7: Resultados LBP y LDP base de datos gb2s, Subregiones de 8x8, radio 1 y 8 vecinos

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Chi-Square)**
LBP Subregiones=8x8 Radio=1 Vecinos=8	70	10	9,29 %
LDP Subregiones=8x8 Radio=1 Vecinos=8	70	10	12,36 %

Experimento base de datos propietaria del grupo gb2s

Del mismo modo que en el experimento anterior, por consistencia con la evaluación de la aproximación geométrica y los resultados previos, se ha utilizado la misma distribución de imágenes y variado del mismo modo el tamaño de las mismas. Como uno de los problemas de la anterior base de datos podría ser el tamaño de las imágenes, se ha vuelto a optar por usar resoluciones mayores para la imagen de la cara. Se ha usado en primer lugar un tamaño de 128x128 píxeles.

En este caso, se aplicarán los dos métodos de textura explicados en el capítulo 3, LBP y LDP. Al presentar cierta variabilidad pero no tan extrema como el caso anterior (FaceScrub), puede ser una buena base de datos para comparar los resultados de estos dos métodos. Los resultados obtenidos por ambos métodos se muestran en la Tabla 4.7.

Los resultados muestran una capacidad de reconocimiento mucho mayor que en el caso del experimento anterior. En este caso la tasa de reconocimiento correcto se encuentra en torno al 90 % para ambos métodos entre si. Se puede ver que los resultados obtenidos con LBP son superiores a los de LDP, por lo que nos centraremos en este método en las siguientes pruebas.

Con el objetivo de encontrar la mejor configuración de parámetros para el método, se ha realizado una exhaustiva batería de pruebas utilizando distintos valores para el número de vecinos, tamaño de las subregiones y el radio. Todas las pruebas realizadas se pueden encontrar en el Anexo B.1. Para este estudio se mantiene la misima división de muestras que en experimentos anteriores (70 % de entrenamiento y 30 % de acceso). Además, se han utilizado 2 tamaños de imágenes: 128x128 y 500x5000, para poder evaluar la influencia del tamaño de imagen. Tras realizar todo este estudio se obtienen resultados realmente positivos en muchas de estas configuraciones. En el tamaño 128x128 la mejor configuración es aquella que tiene 16 vecinos, con radio 2 y tamaño de subregiones 16x16, que ofrece un EER del 5,04 %. Además, se observa que los mejores resultados se obtienen cuando el tamaño de las subregiones es de 16x16, de lo que se puede deducir que no es recomendable submuestrear la imagen en tamaños muy pequeños.

Respecto a las pruebas realizadas con imágenes de tamaño 500x500, los resultados son todavía mejores. El mejor resultado se obtiene para la siguiente configuración: 16 vecinos, con radio 1 y tamaño de subregiones 32x32. Esta configuración devuelve un EER del 0,22 %, es decir, una probabilidad de 99.78 % de verificar al usuario correctamente. Es un gran resultado en el que influye tanto el tamaño de imagen como las características de la propia base de datos. Además se puede advertir que las mejores configuraciones son aquellas que usan subregiones de 32x32, seguidas de las que usan un tamaño de 16x16.

Experimento base de datos BioID

Al igual que en las pruebas de los puntos característicos faciales, en esta base de datos nos centraremos en imágenes 128x128 debido al tamaño inicial de las mismas.

Siguiendo el modelo de estudio de la anterior base de datos, se ha hecho otra exhaustiva batería de pruebas para poder encontrar la mejor configuración de parámetros del método LBP para esta base de datos. Se ha vuelto a analizar la configuración de los parámetros: número de vecinos, tamaño de las subregiones y el radio. Manteniendo la división 70/30 del conjunto de datos en entrenamiento y test. El mejor resultado en términos de EER se ha obtenido para la configuración: 8 vecinos con radio 1 y tamaño de las subregiones 8x8. El valor del EER es de 5.47 %, por lo que el 95 % de los intentos, verificaría al usuario correctamente. En el Anexo B.2 se pueden observar todos los resultados ofrecidos por esta batería de pruebas.

4.3.3. Fusión Geométrica - Holística (LBP)

Tras evaluar los métodos geométrico y holísticos, se ha realizado la evaluación de la fusión de ambas aproximaciones. Para esta evaluación se ha utilizado la base de datos BioID, ya que es una base de datos creada en entornos no controlados y ha dado unos buenos resultados en ámbos métodos. Para la fusión se han utilizado las configuraciones de parámetros que ofrecieron mejores resultados durante la evaluación individual:

- *Landmarks*: DLIB con 47 distancias y 18 ángulos
- Textura: LBP con 8 vecinos, radio 1 y tamaño de las subregiones 8x8.

Los resultados de la fusión se muestran en la Tabla 4.8. Tal y como se explicó en la Sección 3.4, la fusión ha sido realizada utilizando cuatro reglas ponderadas: mínimo, máximo, suma y producto, y precedida de la normalización *min-max* de los *scores*.

En los resultados obtenidos se puede observar que los valores para las diferentes reglas varían bastante. Aun así, todos estos valores mejoran el valor de EER que devolvía la mejor configuración geométrica, que era de 11.87 %. Respecto al mejor valor devuelto por la aproximación holística utilizando LBPs ,5,47 %, solamente la regla del mínimo lo mejora mientras que la regla de la suma prácticamente lo iguala. Sin embargo, si bien se aprecia una cierta mejora en el caso de la fusión con la regla del mínimo, ésta no es lo suficientemente grande como para justificar los cómputos adicionales requeridos para el cálculo de dos métodos y su fusión.

Tabla 4.8: Resultados de la fusión para la base de datos BioID

Regla de normalización	**Regla de fusión**	**EER**
Min - Max	Mínimo	5,28 %
	Máximo	7,08 %
	Producto	8,90 %
	Suma	5,48 %

5

Conclusiones y trabajo futuro

5.1. Conclusiones

En este Trabajo de Fin de Máster se han abordado el estudio, el análisis, el desarrollo la implementación y la evaluación de un sistema biométrico completo facial aplicable tanto en entornos controlados como no controlados. El sistema incluye diferentes técnicas que utilizan la textura y la geometría de la cara como elementos distintivos de los individuos. A pesar de los múltiples trabajos y estudios relacionados con el reconocimiento facial existentes hasta la fecha, por el momento no existe demasiada documentación acerca del reconocimiento facial basado en patrones geométricos. Es por ello que se han utilizado distintos métodos para la extracción de los puntos característicos de la cara. Además, se han utilizado descriptores de textura bien conocidos para extraer los patrones holísticos, más fiables y sencillos. Estos métodos han sido evaluados individualmente utilizando diversas bases de datos. Además, se ha tratado de mejorar las tasas de verificación correcta de los usuarios mediante la fusión de información proveniente de ambas aproximaciones.

Del análisis de los resultados se pueden extraer las siguientes conclusiones:

1. El tamaño de la imagen es muy importante. De las pruebas realizadas con diferentes tamaños de imagen se desprende que un tamaño mínimo de 500x500 es recomendable para obtener resultados aceptables, debido a que la cantidad de información contenida en las mismas es mayor. Una vez superado este tamaño, no se observa una mejora significativa.

2. Las características de las imágenes utilizadas en la evaluación influencian directamente en el resultado. Cuanto menos variabilidad respecto a la apariencia o la expresión de los usuarios, mejores resultados se obtienen. Al estar orientado a dispositivos móviles, el entorno en el que se opere el sistema será semi-controlado. Las personas tienden a usar el teléfono móvil siempre en la misma posición de respecto a la cara, por lo que los elementos que variarán con más frecuencia serán las condiciones de iluminación y fondo, en relación a las expresiones y apariencia de los usuarios.

3. De los resultados ofrecidos por la aproximación geométrica se puede deducir que:

   - La cantidad y precisión de los puntos extraídos es fundamental. Como se ha observado en los resultados, la aproximación que utiliza la API de Google que extrae 8 *landmarks*, tiene unos resultados mucho más deficientes que aquella que utiliza DLIB, que extrae un mayor número de *landmarks* y mucho más precisos.
   - El mayor número de *landmarks* y el mayor número de distancias no hace que los resultados sean mejores. Como se ha podido ver, existen puntos superfluos entre los que apenas existe separación y, por tanto, distancias muy pequeñas como para que exista variabilidad entre distintos usuarios. Por tanto, extraer los *landmarks* y las distancias más representativas de un rostro es una tarea crítica para obtener mejores resultados.

4. Respecto a los resultados derivados de los métodos de textura, las conclusiones son las siguientes:

   - Se han usado dos métodos basados en texturas: LBP y LDP. Al contrario de lo esperado, los resultados reflejan que el método LBP es más eficiente y consigue mejores resultados que el método LDP.
   - Finalmente, en el estudio de la configuración de parámetros del método LBP, se ha observado que no existe una configuración óptima para todos los casos. Si es cierto, que los resultados tienden a ser mejores cuando el tamaño de las subregiones son más pequeñas, debido a que se captura mayor información global, pero que submuestrear demasiado puede introducir ruido. También puede que el número de vecinos tenga influencia siendo los resultados con 16 vecinos mejores que con 8 en la mayoría de los casos. Respecto al radio no se han obtenido datos concluyentes.

5. En cuanto a la fusión, se han extraído las siguientes conclusiones:

   - La fusión mejora las prestaciones del método geométrico sin depender de la regla utilizada. Mientras que para el método holístico utilizando LBPs solo el mínimo lo mejora.
   - La mejora obtenida no es lo suficientemente significativa como para justificar el incremento computacional.

5.2. Trabajo futuro

A continuación se muestran algunas posibles líneas con las que continuar la investigación de este TFM:

- Evaluación del sistema con más bases de datos para poder asegurar que las conclusiones extraídas previamente son correctas y poder seguir trabajando sobre las configuraciones de parámetros.
- Creación de una aplicación móvil que integre el sistema completo. Así, será posible testear estos métodos en condiciones reales.
- Evaluación de otras librerías basadas en *landmarks*, o creación de métodos propios que permitan mejorar la precisión en la detección de estos puntos y, por lo tanto, hacer el sistema más fiable y robusto.
- Estudiar otros métodos para el módulo de comparación, como las redes neuronales o las máquinas de vector soporte (SVM).
- Replicar los experimentos de fusión con otras bases de datos para observar si las conclusiones extraídas con la base de datos BioID son correctas.
- Implementar otro tipo de normalización diferente para evaluar su influencia y si es posible mejorar las tasas obtenidas en este trabajo.

Glosario de acrónimos

- **CCA**: Canonical Correlation Analysis
- **C-APCDA**: Cascade Asymmetric Principal Component Discriminant Analysis
- **DCT**: Discrete Cosine Transform
- **DR**: Detection Rate
- **EER**: Equal Error Rate
- **EFC**: Enhanced Fisher Classifier
- **EFM**: Enhanced Fisher Linear Discriminant Model
- **EP**: Energy Probability
- **FAR**: False Accept Rate
- **FFT**: Fast Fourier Transform
- **FNMR**: False Non-Match Rate
- **FMR**: False Match Rate
- **FRR**: False Reject Rate
- **FTA**: Failure to Acquire
- **FTE**: Failure to Enroll
- **GA-ANN**: Genetic Algorithm - Artificial Neural Network
- **HI**: Histogram Intersection
- **LBP**: Local Binary Pattern
- **LBPU**: Local Binary Pattern Uniform
- **LDA**: Linear Discrimant Analysis
- **LDP**: Local Derivative Pattern
- **NN**: Nearest Neighbour
- **PCA**: Principal Component Analysis
- **SVM**: Support Vector Machine
- **UMACE**: Unconstrained Minimum Average Correlation Energy
- **WHT**: Walsh-Hadamard Transform

Bibliografía

[1] R. Henderson F. Deane, K. Barrelle and D. Mahar. Perceived acceptability of biometric security systems, 1995.

[2] S. Massie S. Elliott and M. Sutton. The perception of biometric technology: A survey. *Automatic Identification Advanced Technologies*, pages 259–264, 2007.

[3] N. Kirschnick H. Sieger and S. Moller. User preferences for biometric authentication methods and graded security on mobile phones. *Symposium on Usability, Privacy, and Security (SOUPS)*, 2010.

[4] C. Sánchez-Ávila B. Ríos-Sánchez, M. Viana-Matesanz and J. Guerra-Casanova. Comfort and Security Perception of Biometrics in Mobile Phones with Widespread Sensors. *Group of Biometrics, Biosignals and Security Research Centre for Smart Buildings and Energy Efficiency (CeDInt), Technical University of Madrid*, 2016.

[5] C. Sánchez-Ávila B. Ríos-Sánchez, M. Viana-Matesanz and M. J. Melcón de Giles. A configurable multibiometric system for authentication at different security levels using mobile devices. *Group of Biometrics, Biosignals and Security Research Centre for Smart Buildings and Energy Efficiency (CeDInt), Technical University of Madrid*, 2016.

[6] Rabia Jafri and Hamid R. Arabnia. A Survey of Face Recognition Techniques. *Journal of Information Processing Systems*, 5(2):41–68, June 2009.

[7] Faizan Ahmad, Aaima Najam, and Zeeshan Ahmed. Image-based Face Detection and Recognition: "State of the Art". *International Journal of Computer Science Issues*, 9(6):3–6, 2013.

[8] Tanvi Chauhan and Sunil Sharma. Literature Report on Face Detection with Skin & Reorganization using Genetic Algorithm. *International Journal of Advanced and Innovative Research*, 2(2):256–262, 2013.

[9] Jay Prakash Maurya and Sanjay Sharma. A Survey on Face Recognition Techniques. *Computer Engineering and Intelligent Systems*, 4(6):11–17, 2013.

[10] Mislav Grgic and Kresimir Delac. Face Recognition Homepage. `http://www.face-rec.org/databases/`. Web. Last access December, 2017.

[11] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face Recognition: A Literature Survey. *ACM Computing Surveys (CSUR).*, 35(4):399–458, 2003.

[12] P. Jonathon Phillips, Harry Wechsler, Jeffery Huang, and Patrick J. Rauss. The FERET database and evaluation procedure for face-recognition algorithms. *Image and vision computing*, 16(I 998):295–306, 1997.

[13] P. J. Phillips, H. Moon, S. A. Rizvi, and P. J. Rauss. The FERET evaluation methodology for face-recognition algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(10):1090–1104, 2000.

[14] Duane M. Blackburn, Mike Bone, and P. Jonathon Phillips. Evaluation Report. Technical report, DoD Counterdrug Technology Development Program Office, 2001.

[15] P. Jonathon Phillips, Patrick Grother, Ross J. Micheals, Duane M. Blackburn, Elham Tabassi, and Mike Bone. Face Recognition Vendor Test 2002. Technical Report March, DARPA, 2003.

[16] K. Messer, J. Matas, J. Kittler, J. Luettin, and G. Maitre. XM2VTSDB: The Extended M2VTS Database. *Second International Conference on Audio and Video-based Biometric Person Authentication*, pages 72–77, 1999.

[17] Sébastien Marcel, Christopher McCool, Pavel Matejka, Timo Ahonen, Jan Cernocký, Shayok Chakraborty, Vineeth Balasubramanian, Sethuraman Panchanathan, Chi Ho Chan, Josef Kittle, Norman Poh, Benoît Fauve, Ondrej Glembek, Oldrich Plehot, Zdenek Jancik, Anthony Larcher, Chistophe Lévy, Driss Matrouf, Jean-François Bonastre, Ping-Han Lee, Jui-Yu Hung, Si-Wei Wu, Yi-Ping Hung, Lukás Machlica, John Mason, Sandra Mau, Conrad Sanderson, David Monzo, Antonio Albiol, Hieu v. Nguyen, Li Bai, Yan Wang, Matti Niskanen, Markus Turtinen, Juan Arturo Nolazco-Flores, Leibny Paola García-Perera, Roberto Aceves-López, Mauricio Villegas, and Roberto Paredes. On the Results of the First Mobile Biometry (MOBIO) Face and Speaker Verification Evaluation. *Recognizing Patterns in Signals, Speech, Images and Videos*, pages 210–225, 2010.

[18] A. C. Morris, J. Koreman, H. Sellahewa, J. Ehlers, S. Jassim, and L. Allano. The SecurePhone PDA Database, Experimental Protocol and Automatic Test Procedure for Multimodal User Authentication. Technical report, Saarland University, Institute of Phonetics., 2006.

[19] Aleix Martínez and Robert Benavente. The AR Face Database. Technical report, 1999.

[20] Wen Gao, Bo Cao, Shiguang Shan, Xilin Chen, Delong Zhou, Xiaohua Zhang, and Debin Zhao. The CAS-PEAL Large-Scale Chinese Face Database and Baseline Evaluations. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, 38(1):149–161, January 2008.

[21] Libor Spacek. Face Recognition Data. `http://cswww.essex.ac.uk/mv/allfaces/index.html`. Web. Last access December, 2017.

[22] Mislav Grgic, Kresimir Delac, and Sonja Grgic. SCface âĂŞ surveillance cameras face database. *Multimedia Tools and Applications*, 51(3):863–879, October 2009.

[23] Athinodoros S. Georghiades, Peter N. Belhumeur, and David J. Kriegman. From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660, 2001.

[24] Ralph Gross, Iain Matthews, Jeff Cohn, Takeo Kanade, and Simon Baker. Multi-PIE. *Proceedings of the International Conference on Automatic Face and Gesture Recognition*, 28(5):807–813, May 2010.

[25] Rodney Goh, Lihao Liu, Xiaoming Liu, and Tsuhan Chen. The CMU Face In Action (FIA) Database. *Analysis and Modelling of Faces and Gestures*, pages 255–263, 2005.

[26] Jennifer Huang, Bernd Heisele, and Volker Blanz. Component-Based Face Recognition with 3D Morphable Models. *First IEEE Workshop on Face Processing in Video, Washington, D.C.*, pages 27–34, 2004.

[27] MorphoTrust USA (Safran). MorphoTrust USA. `http://www.morphotrust.com/`. Web. Last access December, 2017.

[28] FaceKey. Biometric Access Control. `http://www.facekey.com/`. Web. Last access December, 2017.

[29] Cognitec. FaceVACS VideoScan - Cognitec. `http://www.cognitec-systems.de/facevacs-videoscan.html`. Web. Last access December, 2017.

[30] ImageWare Systems Inc. GoMobile Interactive. `http://www.iwsinc.com/`. Web. Last access December, 2017.

[31] BioID AG. MyBioID personal recognition: easy, secure online login and identity management. `https://mybioid.com/`. Web. Last access December, 2017.

[32] Biometrica Systems. Focal Point. `http://biometrica.com/focal_point/`. Web. Last access December, 2017.

[33] ITC-irst. SpotIt! `http://spotit.fbk.eu/SpotIt.html`. Web. Last access December, 2017.

[34] ComputerWorld. What is face id apples new facial recognition. `https://www.computerworld.com/article/3235140/apple-ios/what-is-face-id-apples-new-facial-recognition-tech-explained.html`. Web. Last access January, 2018.

[35] TheWeek. How facial recognition technology creeping into life. `http://theweek.com/articles/737750/how-facial-recognition-technology-creeping-into-life`. Web.

[36] Visidon. Visidon. `http://www.visidon.fi/en/Home`. Web. Last access December, 2017.

[37] Technorms. Visidon facial recognition lock apps. `https://www.technorms.com/12163/visidon-facial-recognition-lock-apps`. Web.

[38] Cha Zhang and Zhengyou Zhang. A Survey of Recent Advances in Face Detection. Technical Report June, 2010.

[39] Roberto Brunelli and Tomaso Poggio. Face Recognition: Features versus Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(10):1042–1052, 1993.

[40] Erik Hjelmå s and Boon Kee Low. Face Detection: A Survey. *Computer Vision and Image Understanding*, 83(3):236–274, September 2001.

[41] Saeed Dabbaghchian, Masoumeh P. Ghaemmaghami, and Ali Aghagolzadeh. Feature extraction using discrete cosine transform and discrimination power analysis with a face recognition technology. *Pattern Recognition*, 43(4):1431–1440, April 2010.

[42] Matthew A Turk and Alex P Pentland. Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91, IEEE Computer Society Conference on*, pages 586–591. IEEE, 1991.

[43] Miguel F. Arriaga-Gómez, Ignacio Mendizábal-Vázquez, Rodrigo Ros-Gómez, and Carmen Sánchez-Ávila. A Comparative Survey on Supervised Classifiers for Face Recognition. In Fabio Garzia, Gordon Thomas, and Daniel A. Pritchard, editors, *IEEE 48th International Carnahan Conference on Security Technology*, pages 407–412, Rome, 2014. IEEE.

[44] FacePhi. Face recognition, fraud prevention. `https://www.facephi.com/en/content/banks`. Web.

[45] FaceFirst. Face recognition for banking. `https://www.facefirst.com/industry/banking`. Web.

[46] HowStuffWorks. When your face is your boarding pass. `https://science.howstuffworks.com/transport/flight/modern/when-face-is-boarding-pass.htm`. Web.

[47] Jie Yang, Xilin Chen, and W. Kunz. A PDA-based face recognition system. *Sixth IEEE Workshop on Applications of Computer Vision, Proceedings.*, pages 19–23, 2002.

[48] A. Hadid, J. Y. Heikkilä, O. Silven, and M. Pietikdinen. Face and eye detection for person authentication in mobile phones. *First ACM/IEEE International Conference on Distributed Smart Cameras*, pages 101–108, 2007.

[49] Stavros Paschalakis and Miroslaw Bober. Real-time face detection and tracking for mobile videoconferencing. *Real-Time Imaging*, 10(2):81–94, April 2004.

[50] Paul Viola and Michael Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1:511–518, 2001.

[51] Sung-Uk Jung, Yun-Su Chung, Jang-Hee Yoo, and Ki-young Moon. Real-Time Face Verification for Mobile Platforms. In *Advances in visual computing*, pages 823–832. Springer Berlin Heidelberg, 2008.

[52] Jianfeng Ren, Nasser Kehtarnavaz, and Lorenzo Estevez. Real-Time Optimization of Viola-Jones Face Detection for Mobile Platforms. *Circuits and Systems Workshop: System-on-Chip-Design, Applications, Integration, and Software, IEEE Dallas*, pages 1–4, 2008.

[53] Chee Kiat Ng, Marios Savvides, and Pradeep K. Khosla. Real-Time Face Verification System on a Cell-Phone using Advanced Correlation Filters. *Fourth IEEE Workshop on Automatic Identification Advanced Technologies*, 2005.

[54] Song-Yi Han, Hyun-Ae Park, Dal-Ho Cho, Kang Ryoung Park, and Sangyoun Lee. Face Recognition Based on Near-Infrared Light Using Mobile Phone. In Springer Berlin Heidelberg, editor, *Adaptive and Natural Computing Algorithms*, pages 440–448, 2007.

[55] Qian Tao and Raymond N. J. Veldhuis. Biometric Authentication for Mobile Personal Device. In *Mobile and Ubiquitous Systems - Workshops, 2006. 3rd Annual International Conference on*, pages 1–3, 2006.

[56] Marcos Faundez-Zanuy, Josep Roure, Virginia Espinosa-Duró, and Juan Antonio Ortega. An efficient face verification method in a transformed domain. *Pattern Recognition Letters*, 28(7):854–858, May 2007.

[57] M. Rahman, J. Ren, and N. Kehtarnavaz. Real-time implementation of robust face detection on mobile platforms. In *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 1353–1356, April 2009.

[58] Guillaume Dave, Xing Chao, and Kishore Sriadibhatla. Face Recognition in Mobile Phones. *Department of Electrical Engineering Stanford University, USA*, 2010.

[59] Jianfeng Ren, Xudong Jiang, and Junsong Yuan. A complete and fully automated face verification system on mobile devices. *Pattern Recognition*, 46(1):45–56, January 2013.

[60] M. Faundez-Zanuy. Data Fusion in Biometrics. *IEEE Aerospace and Electronic Systems Magazine*, 20:34–38, 2005.

[61] N. Poh and S. Bengio. Database, protocols and tools for evaluating score-level fusion algorithms in biometric authentication. *Pattern Recognition*, 39(2):223–233, 2006.

[62] K. Nandakumar A. Ross and A. Jain. Handbook of Multibiometrics (International Series on Biometrics). *Secaucus, NJ, USA: Springer- Verlag New York, Inc.*, page 80, 2006.

[63] A. Hadid M. Pietikinen-P. Matejka J. Cernock N. Poh J. Kittler A. Larcher C. Lvy D. Matrouf J. F. Bonastre P. Tresadern C. McCool, S. Marcel and T. Cootes. Bi-modal person recognition on a mobile phone: Using mobile phone data. *Multimedia and Expo Workshops (ICMEW), 2012 IEEE International Conference*, page 635âĂŞ640, 2012.

[64] K. W. Chung D. J. Kim and K. S. Hong. Person authentication using face, teeth and voice modalities for mobile device security. *IEEE Transactions on Consumer Electronics*, 56(4):2678–2685, 2010.

[65] B. Kang and K. Park. A new multi-unit iris authentication based on quality assessment and score level fusion for mobile phones. *Machine Vision and Applications*, 21(4):541–553, 2010.

[66] Z. Liu and C. Liu. Fusion of color, local spatial and global frequency information for face recognition. *Department of Computer Science, New Jersey Institute of Technology, Newark, NJ 07102, USA*, 2010.

[67] Y. Liu P.A. Heng Q.S. Sun, S.G. Zeng and D.S. Xia. A new method of feature fusion and its application in image recognition. *Department of Computer Science, Nanjing University of Science Technology, Nanjing*, 2005.

[68] C. Liu and H. Wechsler. A shape- and texture-based enhanced Fisher classifier for face recognition. *Department of Mathematics and Computer Science, University of Missouri, St. Louis, MO 63121, United States*, 2001.

[69] M. Pietik ÌĹainen T. Ojala and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition, 29(1):*, pages 51–59, 1996.

[70] M. Pietik ÌĹainen T. Ojala and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on, 24(7):*, pages 971–987, 2002.

[71] Vision and Interaction Group. A dataset with over 100,000 face images of 530 people. `http://vintage.winklerbros.net/facescrub.html`.

[72] BioID. The bioid face database. `https://www.bioid.com/facedb/`.

A

Resultados Landmarks con distancias Manhattan y Chebyshev

A.1. Base de datos Face-Scrub

Tabla A.1: Resultados para la base de datos Face-Scrub con imágenes de tamaño 69x69

Método de extracción de características	**%imágenes entrenamiento**	**%imágenes test**	**EER (Manhattan)**	**EER (Chebyshev)**
ANDROID 8 puntos 28 distancias	70	30	37,69 %	37,52 %
DLIB 68 puntos 2346 distancias	70	30	42,07 %	40,20 %
DLIB 18 puntos 18 distancias	70	30	29,19 %	25,89 %

A.2. Base de datos propietaria del grupo gb2s

Tabla A.2: Resultados para la base de datos del grupo gb2s con imágenes de tamaño 128x128

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Manhattan)**	**EER (Chebyshev)**
ANDROID 8 points 28 distances	70	30	33,25 %	31,46 %
DLIB 68 points 2346 distances	70	30	18,79 %	19,56 %
DLIB 18 points 37 distances	70	30	20,18 %	20,58 %
DLIB 18 points 37 distances 18 ángulos	70	30	18,55 %	18,94 %
DLIB 18 points 47 distances 18 ángulos	70	30	18,54 %	18,81 %

Tabla A.3: Resultados para la base de datos del grupo gb2s con imágenes de tamaño 500x500

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Manhattan)**	**EER (Chebyshev)**
ANDROID 8 points 28 distances	70	30	11,11 %	11,03 %
DLIB 68 points 2346 distances	70	30	7,13 %	7,18 %
DLIB 18 points 37 distances	70	30	6,27 %	5,89 %
DLIB 18 points 37 distances 18 ángulos	70	30	6,09 %	5,71 %
DLIB 18 points 47 distances 18 ángulos	70	30	5,80 %	5,51 %

Tabla A.4: Resultados para la base de datos del grupo gb2s con imágenes de tamaño 890x890

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Manhattan)**	**EER (Chebyshev)**
ANDROID 8 points 28 distances	70	30	11,28 %	10,96 %
DLIB 68 points 2346 distances	70	30	7,02 %	7,20 %
DLIB 18 points 37 distances	70	30	6,13 %	5,81 %
DLIB 18 points 37 distances 18 ángulos	70	30	6,09 %	5,72 %
DLIB 18 points 47 distances 18 ángulos	70	30	5,77 %	5,41 %

A.3. Base de datos BioID

Tabla A.5: Resultados para la base de datos BioID con imágenes de tamaño 128x128

Método de extracción de características	**% imágenes entrenamiento**	**% imágenes test**	**EER (Manhattan)**	**EER (Chebyshev)**
ANDROID 8 points 28 distances	70	30	27,88 %	25,68 %
DLIB 68 points 2346 distances	70	30	15,68 %	14,07 %
DLIB 18 points 37 distances	70	30	16,60 %	15,75 %
DLIB 18 points 37 distances 18 ángulos	70	30	12,98 %	12,79 %
DLIB 18 points 47 distances 18 ángulos	70	30	12,57 %	12,11 %

B

Resultados Texturas

B.1. Base de datos propietaria del grupo gb2s

Tabla B.1: Estudio de la base de datos gb2s con el método LBP y tamaño de imagen 128x128

Base de datos	**Tamaño de la subregión**	**Radio**	**Número de vecinos**	**EER (Chi-Square)**
Base de datos del grupo gb2s *Tamaño: 128x128* *Método: LBP*	8	1	8	7,56 %
	8	1	16	8,22 %
	8	2	8	7,87 %
	8	2	16	6,85 %
	8	3	8	11,76 %
	8	3	16	14,71 %
	16	1	8	6,72 %
	16	1	16	6,90 %
	16	2	8	9,30 %
	16	2	16	5,04 %
	16	3	8	5,88 %
	16	3	16	5,46 %
	32	1	8	9,66 %
	32	1	16	9,66 %
	32	2	8	8,40 %
	32	2	16	7,56 %
	32	3	8	7,14 %
	32	3	16	5,75 %
	64	1	8	13,45 %
	64	1	16	12,23 %
	64	2	8	12,61 %
	64	2	16	10,08 %
	64	3	8	11,59 %
	64	3	16	9,24 %

Tabla B.2: Estudio de la base de datos gb2s con el método LBP y tamaño de imagen 500x500

Base de datos	**Tamaño de la subregión**	**Radio**	**Número de vecinos**	**EER (Chi-Square)**
Base de datos del grupo gb2s *Tamaño: 500x500* *Método: LBP*	8	1	8	1,05 %
	8	1	16	0,99 %
	8	2	8	1,18 %
	8	2	16	1,25 %
	8	3	8	1,31 %
	8	3	16	2,97 %
	16	1	8	0,80 %
	16	1	16	0,43 %
	16	2	8	0,92 %
	16	2	16	0,50 %
	16	3	8	1,07 %
	16	3	16	0,87 %
	32	1	8	0,75 %
	32	1	16	0,22 %
	32	2	8	0,70 %
	32	2	16	0,25 %
	32	3	8	0,75 %
	32	3	16	0,31 %
	64	1	8	1,59 %
	64	1	16	1,06 %
	64	2	8	1,22 %
	64	2	16	0,76 %
	64	3	8	0,92 %
	64	3	16	0,61 %

B.2. Base de datos BioID

Tabla B.3: Estudio de la base de datos BioID con el método LBP y tamaño de imagen 128x128

Base de datos	**Tamaño de la subregión**	**Radio**	**Número de vecinos**	**EER (Chi-Square)**
Base de datos BioID *Tamaño: 128x128* *Método: LBP*	8	1	8	5,47 %
	8	1	16	6,85 %
	8	2	8	6,29 %
	8	2	16	9,07 %
	8	3	8	6,86 %
	8	3	16	15,44 %
	16	1	8	8,38 %
	16	1	16	6,13 %
	16	2	8	6,98 %
	16	2	16	6,16 %
	16	3	8	6,76 %
	16	3	16	7,82 %
	32	1	8	10,36 %
	32	1	16	8,81 %
	32	2	8	10,96 %
	32	2	16	9,59 %
	32	3	8	10,62 %
	32	3	16	10,36 %
	64	1	8	10,02 %
	64	1	16	9,89 %
	64	2	8	10,57 %
	64	2	16	10,73 %
	64	3	8	10,13 %
	64	3	16	10,96 %